CODEN: JASI

The Journal of the Acoustical Society of America

0001-4966

Vol. 108, No. 5, Pt. 1 November 2000

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 108, NO. 5, PT. 1, NOVEMBER 2000

CONTENTS—*Continued from preceding page*

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 108, NO. 5, PT. 1, NOVEMBER 2000

CONTENTS—*Continued from preceding page*

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 108, NO. 5, PT. 1, NOVEMBER 2000

CONTENTS—*Continued from preceding page*

LETTERS TO THE EDITOR

ERRATA

CUMULATIVE AUTHOR INDEX

NOTES CONCERNING ARTICLE ABSTRACTS

1. The number following the abstract copyright notice is a Publisher Item Identifier (PII) code that provides a unique and concise identification of each individual published document. This PII number should be included in all document delivery requests for copies of the article.

2. PACS numbers are for subject classification and indexing. See June and December issues for detailed listing of acoustical classes and subclasses.

3. The initials in brackets following the PACS numbers are the initials of the JASA Associate Editor who accepted the paper for publication.

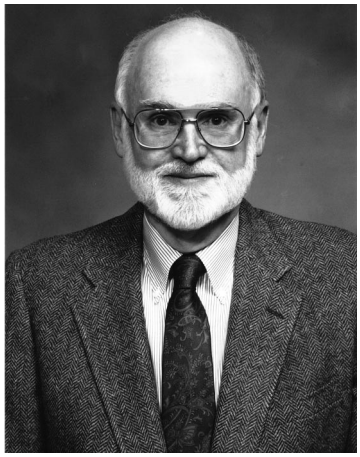
CONTENTS

ACOUSTICAL NEWS—USA

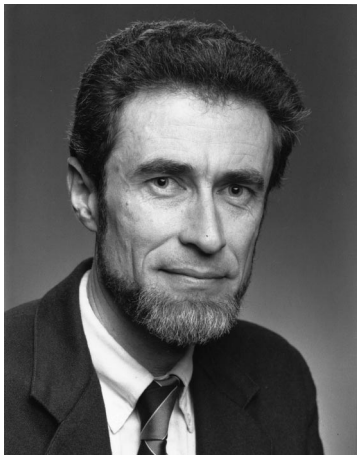
Elaine Moran
Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502

Editor's Note: *Readers of this Journal are encouraged to submit news items on awards, appointments, and other activities about themselves or their colleagues. Deadline dates for news items and notices are 2 months prior to publication.*

New Fellows of the Acoustical Society of America



R. Dean Ayers—For contributions to the understanding of wind instruments and education.



Donald E. Hall—For contributions to musical acoustics and education.

Ji-qing Wang (Chi-ching Wong)—For contributions to architectural acoustics and international collaboration.

Thomas D. Rossing receives the 2000 Robert A. Millikan Medal

Thomas D. Rossing was awarded the Robert A. Millikan Medal of the American Association of Physics Teachers (AAPT) in August 2000 at its Summer Meeting in Guelph, Ontario, Canada. The Millikan Medal is awarded annually to a university or high school teacher nationwide for notable and creative contributions to the teaching of physics. Dr. Rossing's Millikan Medal address was titled ''Beauty in Physics and the Arts.''

Thomas Rossing is Professor of Physics at Northern Illinois University where he established a unique research program in musical acoustics. He has authored more than 300 scholarly papers and 14 books, including *The Science of Sound.* Dr. Rossing has been active in organizing national, state, and regional meetings of physics teachers and is a past president of the AAPT. In addition to studying music, Rossing is a musician. He plays the clarinet and recorder and has also played in the Kishwaukee Symphony in DeKalb as well as singing with the Chicago Symphony Chorus, the San Francisco Symphony Orchestra, the Minnesota Orchestra, and with choruses in Germany, England, Sweden, and Australia.

Tom Rossing is a Fellow of the Acoustical Society of America (ASA) and received the ASA Silver Medal in Musical Acoustics in 1992 ''for major influence on research and teaching in musical acoustics and contributions to the understanding of percussion instruments.'' He served as Chair of the ASA Technical Committee on Musical Acoustics from 1984–1990 and is currently Editor of *Echoes*, the newsletter of the Acoustical Society of America.

Christopher Jaffe honored by the American Institute of Architects

ASA Fellow Christopher Jaffe is one of seven recipients of The Year 2000 Institute Honors for Collaborative Achievement Award of the American Institute of Architects. The award recognizes distinguished achievements of allied professionals whose achievements have had a beneficial influence on or advanced the architectural profession.

Reviewers of Manuscripts, 1998 and 1999

Each year the Journal endeavors to publish a list of all the persons who reviewed manuscripts during the preceding year. Such a list is a compendium of names supplied by the Associate Editors. Because our peer review system depends strongly on the continuing anonymity of the reviewers, the Journal publishes these names in alphabetical order without identification of the associate editors who provided the names and without identification of the papers they reviewed. The last such published list, for 1997, appeared in the July 1998 issue. Because of the temporary disruption caused by the unexpected death of the previous Editor-in-Chief, Daniel Martin, the 1998 list was not published, and the present Editor-in-Chief apologizes for this. The list below is intended to include the names for both 1998 and 1999.

The primary reason for the publication of the list is to express the Journal's gratitude to its reviewers. Reviewing a paper is often a very time-consuming and demanding task, and the anonymity requirement yields no professional recognition to those who generously provide their time to help the Associate Editors decide which papers should be published and to give

constructive criticisms to the authors. The Journal is justifiably proud of this list, which includes a goodly proportion of all the researchers and eminent authorities in acoustics and related fields.

In a compendium of this length, omissions and errors are inevitable. If anyone notices such, please send the corrections and missing names via e-mail or regular mail to either Elaine Moran (elaine@aip.org) or Allan Pierce (adp@bu.edu).

Abdala, Carolina
Abel, Sharon
Abraham, Douglas
Abramson, Arthur S.
Accorsi, M. L.
Achenbach, J. D.
Ackermann, H.
Adachi, Seiji
Adams, S.
Agullo, Joaquim
Akeroyd, Michael
Alipour, F.
Alku, Paavo
Allard, J. F.
Allen, Jont B.
Allen, Prudence
Alers, George
Altes, Richard
Amir, Noam
Anderson, Michael J.
Anderson, Aubrey
Apfel, Robert E.
Arehart, Kathy
Arnott, W. Pat
Arzoumanian, Sevag
Ashmead, Daniel
Askenfelt, Anders G.
Assmann, Peter
Atalla, Noureddine
Atchley, Anthony
Attenborough, Keith
Aubauer, Roland
Audoin, Bertrand
Augspurger, G. L.
Avan, Paul
Ayers, R. Dean
Ayers-Elam, Gayle
Bacon, Sid
Badiey, Mohsen
Badin, Pierre
Baer, Thomas
Baggeroer, Arthur
Bahr, Ruth
Bai, Mingsing R.
Bailey, J. Michael
Bailey, P.
Bailly, Gérard
Bajaj, Anil K.
Baken, R.
Baker, Andrew
Ballato, A.
Barabanenkov, Yuri N.
Barbone, P.
Barbour, Dale
Bard, Benjamin A.
Barjau, Ana
Barney, Anna
Barrett, Harrison H.
Barron, Michael
Bartram, J. F.
Bass, Henry E.
Baum, Shari R.
Beauchamp, James W.
Beckman, Mary Esther
Beddor, Patrice S.
Behne, Dawn
Behrman, Alison
Benjamin, K. C.
Benki, Jose R.
Bennet-Clark, Henry
Bensen, Robert
Beranek, Leo L.
Berg, Bruce
Berg, Kathleen M.
Bernstein, Leslie
Bernstein, Lynne E.
Berry, David
Berthelot, Yves H.
Bilaniuk, N.
Bilsen, Frans A.
Binder, Jeffry
Bishop, Garner C.
Bissinger, George A.
Bistafa, Sylvio
Blackstock, David
Blaurent, Jens
Bloothooft, Gerrit
Boe, L.
Bohn, Ocke-Schwen
Boite, Postale
Bookheimer, Susan
Booth, Newell
Boothroyd, Arthur
Bork, Ingolf
Boulanger, Patrice
Bowles, Ann
Bradley, J. S.
Bradlow, Ann
Braida, Louis D.
Brandstein, Michael
Breazeale, M. A.
Bregman, Al
Briers, Rudy
Briggs, G. A. D.
Briggs, Kevin
Broad, David J.
Bronkhorst, Adelbert
Brown, Ann M.
Brown, Carolyn
Brown, Charles
Brown, Steven M.
Bruggeman, J. C.
Brungart, Douglas
Bucaro, J. A.
Buck, John
Buckingham, Michael
Buder, Eugene H.
Buell, Thomas N.
Bulla, Wesley
Bunnel, H. Timothy
Burgess, William
Burnett, David S.
Burns, Edward M.
Burroughs, Courtney B.
Busby, Peter A.
Busch-Vishniac, Ilene J.
Buss, Emily
Butler, J. L.
Buus, Soren
Byrd, D.
Cacace, Anthony T.
Canlon, Barbara
Carey, William
Cariani, Peter
Carlie, Simon
Carlyon, Robert P.
Carney, Arlene Earley
Carre, R.
Carstens, B.
Carstensen, Ed
Carter, G. Clifford
Carvalho, A. P. O.
Caspers, Johanneke
Castagnede, Bernard
Castaings, Michel
Cato, Doug
Causse, Rene E.
Cavanagh, Ray
Cawley, Peter
Ceperley, P. H.
Chafe, Chris
Chalikian, Tigram V.
Challis, Richard E.
Chambers, James
Champlin, Craig
Chamuel, Jacques R.
Champagne, Benoit
Chapman, David
Chapman, Ross
Chatterjee, Monita
Cheeke, J. N. D.
Cheng, A. C. H.
Cheyne, Stanley A.
Childers, Donald G.
Chimenti, Dale E.
Chinnery, Paul A.
Chotiro, Nicholas
Chu, Wing T.
Chu, Y. C. (Ya-Cherng)
Church, Charles C.
Claesson, Ingvar
Clark, Christopher
Clarke, Eric F.
Clarkson, Marsha
Clay, Clarence
Cleveland, Robin
Clifton, Rachel
Cohen, Michael
Colburn, Steven H.
Collet, Lionel
Collins, Leslie M.
Collins, Michael D.
Coltman, John W.
Colton, Raymond H.
Conklin, Harold A., Jr.
Constable, T.
Cook, Perry R.
Coomb, Sheryl
Copley, David C.
Cops, Andre
Costley, R. Daniel
Cottingham, James P.
Craig, Jim
Craik, R. J. M.
Cranen, B.
Creamer, Dennis B.
Crum, Lawrence
Culick, Fred E. C.
Culling, John
Cummings, William
Cummins, Fred
Cunefare, Kenneth A.
Cuomo, F. W.
Cuschieri, J. M.
Dahl, Peter
Dai, Huanping
Daigle, G. A.
Dalsgaard, Jakob
D'Amico, Angela
Dang, Jianwu
Daniloff, R.
Dannenberg, Roger
Darling, Timothy
Darvennes, Corinne
Darwin, Christopher J.
Dau, Torsten
Davies, W. J.
Davis, Evan B.
Davis, Katherine
Dayal, Vinay
Deane, Grant B.
de Billy, Michel E.
de Cheveigne, Alain
DeJong, Kenneth
de Krom, Guus
Delgutte, Bertram
Demany, Laurent
Denardo, Bruce
Deng, Li
de Nil, L.
Deschamps, Marc
Desmet, W.
Deutsch, Diana
Dickey, Joseph
Diehl, Randy
Dillon, Harvey A.
Ditri, John J.
Divenyi, Pierre
Dizon, Bob
Dolan, David F.
Dolphin, William
Donaldson, Gail S.
Dooley, Gary
Dooling, Robert J.
Dorfman, Robert J.
Dorman, Michael F.
Dorn, Patricia
Dosso, Stan
Dowling, Ann
Dowling, David R.
Downing, J. Micah
Dreschler, Wouter
Drumheller, Douglas S.
D'Spain, Gerald
Dubno, Judy R.
Duda, Richard
Duda, Tim
Duifhuis, Hendrikus

Dunn, Floyd
Durlach, Nathaniel I.
Dye, Raymond H.
Eddins, David A.
Edwards, Brent
Edwards, Jan
Egan, M. David
Eggermont, Jos
Egle, David M.
Eilers, R.
Eisenberg, Laurie
Elliott, Stephen
Embleton, T. F. W.
Embrechts, J. J.
Eneroth, Peter
Engebretson, Maynard A.
Epifanio, Chad
Erbe, Christine
Espy-Wilson, Carol Y.
Evans, Richard
Evans, William
Everstine, Gordon C.
Every, Arthur C.
Ewart, Terry E.
Faber, Alice
Fahey, Paul
Fahnline, John B.
Fahy, Frank
Fantini, Deborah A.
Farabee, Theordore M.
Farassat, F.
Fastl, Hugo
Faulkner, Andrew
Fawcett, John A.
Fay, Richard
Feke, Donald L.
Feng, Leping
Ferguson, Brian G.
Ferrigno, G.
Festen, Joost
Feth, Lawrence L.
Fialkowski, Laurie T.
Fiddy, Michael
Fidell, Sanford
Finch, Robert D.
Finette, Steve
Finley, Charles C.
Finneran, James
Fisher, Fred
Fitzgerald, Edwin R.
Fitzgibbons, Peter J.
Flanagan, J. L.
Flandro, Gary A.
Flege, James
Fletcher, Neville
Florentine, Mary
Floyd, Robert
Formby, Craig
Forrest, Karen M.
Forrest, Timothy
Fougeron, Cécile
Fourakis, Marios
Fowlkes, J. Brian
Fox, Christopher
Fox, Robert A.
Francois, Robert
Franke, S. J.
Frankel, Adam
Franzoni, Linda
Frazer, L. Neil
Freyman, Richard L.
Friberg, Anders
Frisk, George
Frison, T.
Fristrup, Kurt
Frizzell, Leon
Fromkin, Victoria A.
Fromm, David
Fu, Qian-Jie
Furukawa, Shigeto
Fuzessery, Zoltan (Nick)
Gabillet, Yannick
Gabrielson, Thomas B.
Gabrielsson, Alf
Gadd, Charles W.
Gade, Anders C. S.
Gaensler, Tomas
Gaitan, D. Felipe
Garcia-Bonito, J.
Gardner, David
Garrelick, Joel
Garrett, Steve
Gaumond, Chares F.
Gaunaurd, Guillermo C.
Gaunt, Sandra
Gauss, Roger
Gautesen, A. K.
Geers, Thomas L.
Gentner, Timothy
Gerratt, Bruce R.
Gerstoft, Peter
Gescheider, George
Gilbert, Joel
Gilbert, Kenneth
Gilkey, Bob
Giordano, Nicholas J.
Gladwell, Graham
Gockel, Hedwig
Goff, John A.
Goldstein, Julius L.
Goldstein, Louis M.
Goodman, Ralph
Gordon, Peter C.
Gorga, Michael P.
Gotteland, Maurice
Gow, Merrilyn
Grabe, E.
Gracco, Vincent
Grace, Sheryl M.
Granato, A. V.
Grant, Kenneth
Gray, Doug
Gray, Stephen
Green, Kerry
Greenberg, Julie
Greenberg, Steven
Greene, Charles
Grennlaw, Charles F., III
Griesinger, D.
Gronnum, N.
Grose, John H.
Gu, Yonghai Chris
Guenther, Frank
Guerich, Mohamed
Guo, Ningqun
Gusev, Vitalyi
Gussenhoven, Carlos
Guy, Richard W.
Guyer, Bob
Guzman, Sandra
Haberman, Robert C.
Hafter, Ervin
Hagelberg, Paul M.
Hagiwara, R.
Hall, Donald E.
Hall, James W., III
Hall, Joseph L.
Hall, Marshall
Hallaj, Ibrahim
Hambric, Steven A.
Han, Mieko S.
Han, Wei
Hanson, Helen M.
Hanson, Roger J.
Hardin, Jay C.
Harnsberger, James D.
Harris, Frances P.
Harris, Gerald R.
Harris, John G.
Hartmann, William M.
Hasegawa, Takahi
Hasegawa-Johnson, M.
Hastings, Mardi
Havelock, David I.
Hayek, Sabih I.
Hazony, Dov
Heald, Gary
Heaney, Kevin D.
Heard, Garry J.
Heffner, Henry
Hefner, Brian T.
Helfer, Karen S.
Heller, Laurie
Hellman, Bill
Hellman, Rhona
Helweg, David
Henyey, Frank S.
Herman, Cila
Hermes, Dik J.
Herstein, L. A.
Hertrich, Ingo
Herzel, Hans-Peter
Herzing, Denise
Heyliger, Paul
Hickey, Craig
Hickling, Robert
Hicks, Michelle L.
Hillenbrand, James M.
Hiller, Robert
Hinich, Melvin J.
Hirao, Masahiko
Hirschberg, Avraham
Hirst, Daniel J.
Hixon, E. L.
Hixon, Thomas J.
Hjelmstad, K. D.
Hlawatsch, Franz
Hodge, Megan
Hodgkiss, William
Hodgson, M. R.
Hoekje, Peter L.
Hoffmann, Thomas L.
Hogan, Sarah
Hogden, John
Hoit, J.
Hole, Lars R.
Holliday, D. Vance
Hollien, Harry
Holt, Lori
Holt, R. Glynn
Honda, Kiyoshi
Hoole, P.
Hornsby, Benjamin
Horwitz, Amy
Hossack, J. A.
Hosten, Bernard
Houtsma, Adrianus
Hovem, Jens M.
Howarth, Thomas R.
Howe, Michael S.
Hoy, Ronald
Hsu, David K.
Hsu, R. M. F.
Huang, Dehua
Huang, Wei
Hukin, Robert
Huron, David
Hutchins, Carleen M.
Hutchins, D. A.
Hutchinson, James R.
Hwang, Yun-Fan
Idogawa, Tohru
Igusa, Takeru
Ih, Young-Sang
Imaizumi, Satoshi
Ingard, K. U.
Ioup, Juliette W.
Irino, Toshio
Isaak, Donald G.
Iverson, Paul E.
Iyer, Nandini
Jackson, Darrell R.
Jackson, M.
Jackson, Michel T. T.
Jacobsen, Finn
Jaeger, Jeri J.
Janik, Vincent
Jansson, Erik V.
Javel, Eric
Jech, J. Michael
Jen, Philip
Jenison, Rick
Jensen, Finn
Jesteadt, Walt
Jesus, Sergio
Jezequel, L.
Jia, X.
Johnson, David L.
Johnson, Keith
Jones, Ken
Jones, D. F.
Jongman, Allard
Junger, M. C.
Kaduchak, Greg
Kaernbach, Christian
Kalinowski, J.
Kaminsky, I.
Kang, S. W.
Kaplunov, J. D.
Kargl, Steven
Karlsson, I.
Karnell, Michael P.

Karnell, M.
Kastak, David
Kates, James M.
Katz, Richard
Katz, W.
Kauffmann, Christiaan
Kawahara, Hideki
Kay, Steven
Keefe, Douglas H.
Keith, William L.
Keller, Joseph B.
Kellermann, Walter L.
Keltie, Richard
Kempster, G.
Keolian, Robert
Keppens, Veerle
Kergomard, Jean
Kerry, Geoff
Kessinger, Rachel H.
Khadar, M. Abdul
Kharakoz, Dmitry
Khuri-Yakub, Butrus
Kibblewhite, Ian
Kidd, Gary R.
Kidd, Gerlad D.
Kiefte, Michael
Kil, Hyun-Gwon
Kilhman, Tor. S. D.
Killion, Mead
Kim, Duck O.
Kim, Kwang Yul
Kim, Yang-Hann
Kimball, Christopher
Kimberley, Barry
Kindel, Joe
Kingston, John
Kirsch, Andreas
Kirstens, Ivars P.
Kistler, Doris
Kleiner, Mendel
Kleppe, John A.
Klosner, Jerome M.
Knobles, David P.
Ko, S. H.
Koehnke, Janet
Koessl, Manfred
Kohlrausch, Armin
Kollmeier, Birger
Krakow, R.
Krane, M.
Kreiman, J.
Kröger, Berndt J.
Krothapalli, A.
Kuc, Roman
Kuehn, David P.
Kuhl, Patricia
Kujawa, Sharon G.
Kumada, Masanobu
Kumar, Arun
Kuperman, William A.
Kuwano, Sonoko
Laboissière, Rafael
Ladefoged, Peter
Lafarge, Denis
Lafleur, L. Dwynn
Lakes, Rod
Laloe, Franck
Lam, Y. W.
Lane, Harlan
Langley, R. S.
Laron, David
Larraza, Andres
Larson, Charles
Lasky, Robert E.
Lauchle, Gerald C.
Lauriks, Walter
Lauterborn, Werner
Laville, Frederic
Lavrentyev, Anton I.
Leclaire, Ph.
Ledbetter, Hassel
Leek, Marjorie
Leep-Apolloni, Laurine
Leissa, Arthur W.
Leisure, Robert
Lenoir, Olivier
LePage, Kevin
Le Prell, Colleen Garbe
Leroy, Oswald
Letcher, Stephen V.
Levesque, Daniel
Levitt, Harry
Levy, Moises
Lewin, Peter A.
Leybourne, Allen E.
Li, Yuang-Liang
Liénard, Jean-Sylvain
Lightfoot, Jay
Liljencrants, Johan C. W. A.
Lim, C. W.
Lim, Raymond
Lin, C. Y.
Lin, Y. K.
Lindblom, B.
Lipkens, Bart
Lisker, Leigh
Liss, J.
Litovsky, Ruth
Litt, Mitchell
Liu, Yijun
Ljunggren, Sten
Llewellyn Smith, Stefan
Lobkis, Oleg I.
Loeffler, C. M.
Logan, Kenneth J.
Logan, K.
Loizou, Philip C.
Lomonosov, Alexei
Long, Glenis R.
Lonsbury-Martin, Brenda
Lotto, Andrew J.
Love, Richard
Lowe, M. J. S.
Loyau, Thierry
Lu, I-Tai
Lu, Yichi
Lucero, Jorge C.
Lung, Alex
Lutfi, Robert
Luz, George
Lynch, James
Lyons, Anthony P.
Lyzenga, Johannes
MacDougall, Scott
Mace, B. R.
MacLennan, David N.
Macpherson, Ewan
Madden, John
Mahshie, J.
Makino, Shoji
Malme, Charles
Mann, David
Mann, J. A., III
Mann, Jakob
Mantena, P. Raju
Margolis, Robert H.
Margulies, Timothy S.
Marques, Wilson, Jr.
Marsh, Alan H.
Marshall, Gerald L.
Marshall, Harold
Marshall, Lynne
Marston, Philip L.
Martin, D. W.
Martin, Glen K.
Martinez, R.
Martinez, Tony R.
Mason, Christine R.
Massaro, Dominic
Masson, Patrice
Mast, T. Douglas
Masters, Mitch
Matsukawa, Mami
Mattei, P. O.
Matthies, Melanie
Mattingly, Ignatius M.
Matula, Thomas
May, Bradford J.
Mayer, Larry
Mayer, W. G.
Maynard, Julian D., Jr.
Maze, Gerard
McAdams, Stephen E.
McCall Denham, Susan
McCammon, Diana
McCoy, John J.
McDaniel, Suzanne
McDaniel, J. G.
McDermott, Hugh
McDonald, Mark
McDonald, Ed
McFadden, Dennis
McFarland, David
McGehee, Duncan
McGowan, Richard S.
McIntyre, Michael E.
McKay, Colette M.
McPherson, Mike
Means, Steve
Mecado, Eduardo, III
Meddis, Raymond
Meegan, G. D.
Mellen, Robert H.
Mellert, Volker
Mellinger, David
Mendor, William E.
Mergell, Patrick
Mershon, Donald
Metzger, Kurt
Meynial, Xavier
Michaelis, Dirk
Michaels, J. E.
Michalopoulou, Zoi-Heleni
Michelsen, Axel
Middlebrooks, John
Migliori, Albert
Mikhalevsky, Peter
Miki, N.
Miller, Eric
Miller, Lee
Miller, James
Miller, James G.
Mills, David M.
Ming, R.
Mobley, Joel
Moffett, M. B.
Mohl, Bertl
Moldover, Michael L.
Moller, Henrik
Molyneux, John E.
Molz, Eric B.
Mongeau, Luc
Moore, Brian C. J.
Moore, Christopher A.
Moore Patrick
Morimoto, M.
Morishima, S.
Morro, Angelo
Moss, Cynthia
Moulin, Annie
Mountain, David
Mozurkewich, George
Muijser, Hans
Muller, Rolf
Munhall, Kevin G.
Munjal, M. L.
Munshi, T. K.
Murphy, Bill
Murphy, Peter
Myers, Michael K.
Myrberg, Arthur
Nabelek, Anna K.
Nachtigall, Paul
Nagaya, K.
Naghshineh, Koorosh
Nagy, Peter B.
Nair, S.
Nakamura, Akira
Narayanan, Shrikanth S.
Narins, Peter
Nayfeh, Adnan H.
Naylor, Graham
Nederveen, Cornelis J.
Neely, Stephen
Neff, Donna Lynn
Nejade, Armand
Nelisse, Hugues
Nelson, David A.
Nelson, Peggy B.
Nero, Redwood W.
Newman, Rochelle
Newton, Roger G.
Nielsen, Johann
Nielsen, Peter
Nightingale, Trevor N. T.
Nittrouer, Susan N.
Noble, William
Nolle, A. Wilson
Norris, A. N.
Norris, Jeffrey
Novarini, Jorge C.
Nozza, Robert J.

Nuttall, Alfred L.
Nye, Patrick W.
Nygaard, Lynne C.
Nystuen, Jeff
O'Brien, William D.
Oda, Hitoshi
Ogi, Hirotsugu
Oguz, Hasan
Ohde, Ralph N.
Ohde, R.
Ohya, Akihsa
Onodera, Y.
Opie, Jane
Orduna-Bustamante, Felipe
Orlikoff, Robert F.
O'Shaughnessy, Douglas D.
Osler, John
Ostry, D.
Otashev, Vladimir E.
Ostrovsky, Lev
Oxenham, Andrew J.
Pace, Nick
Palmer, Alan R.
Palmer, Caroline
Palmer, Catherine V.
Pan, Jiaquiang
Pan, Jie
Panneton, Raymond
Parker, Kevin J.
Parncutt, Richard
Pastore, Richard
Pater, Larry
Patterson, Roy D.
Pavlovic, Chaslav
Payne, Katie
Pecorari, Claudio
Pelorson, Xavier
Perrier, Pascal
Perrott, David R.
Peters, Michael
Peters, Robert W.
Pezerat, Charles
Pfingst, Bryan E.
Phillips, Dennis P.
Photiadis, Douglas M.
Pickering, Norman C.
Pierce, Allan
Pierucci, Mauro
Pimshtein, Valerii G.
Pind, Jörgen
Piquette, Jean C.
Plack, Christopher J.
Plona, Thomas J.
Plotkin, Kenneth J.
Podlesak, Michael
Polka, Linda
Pollard, Howard F.
Poncelet, O.
Ponton, Curtis
Popelka, Gerald R.
Popovics, J. S.
Popper, Arthur
Port, Robert F.
Pouliquen, Eric
Povey, Malcolm
Powell, A.
Powell, Clemens
Prada, Claire
Preston, John R.
Preves, David
Prieve, Beth
Probst, Rudolf
Prospretti, Andrea
Prosser, William H.
Puckette, Miller
Puria, Sunil
Pyle, Robert W., Jr.
Pyne, James M.
Qi, Yingyong
Queheillalt, Douglas T.
Raatgever, Johan
Rabbit, Richard D.
Rabinkin, D. V.
Radebaugh, Ray
Raichel, Daniel R.
Rakerd, Brad
Ramig, L.
Ramsay, Gordon
Rankovic, Christine M.
Rasmussen, Karsten Bo
Raspet, Richard
Ratner, N.
Reed, Charlotte M.
Regazzioni, Carlo
Reinhardt-Rutland, Anthony
Remez, R.
Ren, Tianying
Rhode, William
Richards, Virginia M.
Richardson, Bernard E.
Richardson, John
Richardson, Michael D.
Rickert, Martin
Ricks, David C.
Ridgway, Sam
Rindel, Jens Holger
Rink, John
Risset, Jean-Claude
Roberts, Brian
Roberts, Daniel
Roberts, Gareth W.
Robinson, Donald
Robinson, H. C.
Roch, Maurice
Rodet, Xavier
Rodgers, Oliver E.
Rogers, P. H.
Roitblat, Herbert
Rokhin, S. I.
Rolt, K. D.
Rose, James H.
Rose, Joseph L.
Rosen, Stuart V.
Rosenblum, Lawrence
Rosowski, John J.
Rossing, Thomas D.
Rouseff, Daniel
Roux, Philippe
Roy, Ronald A.
Royston, Thomas
Ruben, Robert
Rubin, Philip E.
Rubinstein, Jay
Rubio, Javier Giner
Ruckman, Christopher D.
Ruiz, Robert
Ryan, James
Sabatier, James M.
Saberi, Kourosh
Sadler, Brian M.
Safaeinili, Ali
Salisbury, John I.
Salvi, Richard
Sammelmann, Gary S.
San Emeterio, Jose Luis
Sandell, Gregory J.
Sandman, Bruce
Sanguineti, Vittorio
Sapienza, Christine H.
Sas, P.
Scandrett, Clyde
Scavone, Gary
Scharf, Robert
Scheirer, Eric
Scherer, Ronald C.
Schlauch, Robert S.
Schmeidt, Richard
Schmerr, Lester W., Jr.
Schmid, Sabina
Schmiedt, Richard
Schneider, John
Schniztler, Hans Ulrich
Schoentgen, Jean
Schroeder, Manfred R.
Schultheiss, Peter
Schumacher, Robert T.
Schwartz, Stuart
Selamet, Ahmet
Sessler, Gerhard M.
Sethares, W. A.
Seybert, Andrew F.
Sgard, Franck
Shadle, C.
Shaffer, L. Henry
Shamma, Shihab
Shang, Er-Chang
Shannon, Robert V.
Sharp, Ben H.
Shaw, Edgar A. G.
Sheft, Stanley E.
Shen, I. Y.
Shera, Christopher
Sherman, Charles H.
Shield, Bridget M.
Shields, F. Douglas
Shimauchi, Suehiro
Shinn-Cunningham, Barbara
Shipp, Thomas
Shoonhaven, Ruurd
Shosani, Herzy
Shuyu, Lin
Sibul, Leon
Siderius, Martin
Siebein, Gary W.
Siegel, Jonathon H.
Simmons, James
Simpson, Harry
Sinder, D.
Singher, Liviu
Sinha, Dipen
Slaney, Malcolm
Smith, Barry T.
Smith, Bruce L.
Smith, David
Smith, Julius O., III
Smith-Olinde, Laura
Smith, Stuart T.
Smoorenburg, Gudio F.
Smurzynski, Jacek
Snell, Karen B.
Soize, Christian
Sommerfeldt, Scott
Sommers, Mitchell
Sondhi, M. M.
Song, Hee Chun
Sorbello, Richard S.
Sorokin, V.
Soulodre, Gilbert A. J.
Souza, Pamela
Sparrow, Victor W.
Spetzler, Hartmut
Spicer, James B.
Spies, M.
Spiesburger, John
Spindel, Jonathan
Spindel, R. C.
Stanton, Tim
Stathopoulos, Elaine T.
Stellmack, Mark
Stelmachowicz, Patricia G.
Stepanishen, Peter R.
Stephen, Ralph
Stern, Richard
Stevens, Kenneth N.
Stinson, Michael R.
Stoel-Gammon, C.
Stoll, Robert
Stone, Maureen L.
Story, Brad H.
Stoughton, Roland
Stover, Lisa J.
Strange, Winifred
Strasberg, Murray
Strickland, Elizabeth A.
Strik, Helmer
Strong, William Y.
Studebaker, Gerald A.
Stumpf, Folden B.
Sullivan, D. L.
Sullivan, E. J.
Summerfield, Quentin
Summers, Ian
Summers, Van
Sundberg, Johan E. F.
Supin, Alexander
Surlykke, Annemarie
Sussman, Joan E.
Sussman, Harvey
Suter, A.
Suzuki, Hideo
Suzuki, Yo iti
Svensson, U. Peter
Svirsky, Mario A.
Swift, Greg W.
Syrdal, A.
Szabo, Thomas L.
Szwerc, Richard
Tague, John
Takahashi, D.
Talaske, R. H.
Talmadge, Carrick L.
Tan Jinxing

Tang, Dajun
Tarnow, Viggo
Taroudakis, Michael
Telschow, Ken
Temkin, Samuel
TenCate, James
Terhardt, Ernst J.
Ternstrom, Sten
Terrill, Eric
Thenail, Denis
Thoen, Jan
Thompson, Rex P.
Thompson, Stephen C.
Thompson, William
Thomson, David J.
Thorne, Peter
Thornton, Roger
Tindle, Christopher T.
Ting, R. Y.
Tittmann, Bernhard R.
Titze, Ingo R.
Tjotta, Sigve
Tobocman, W.
Todd, Neil
Tolin, Daniel
Tolstoy, Alex
Trahiotis, Constantine
Traykovski, Peter
Trevorrow, Mark V.
Troeger, Richard
Tubis, Arnold
Tuller, Betty H.
Turner, Christopher W.
Tyack, Peter
Tyler, Richard S.
Uchanski, Rosalie
Umeda, Yoshikuni
Ungar, Eric E.
Vagle, Svein
Van Buren, Arnie Lee
Vance, Timothy J.
Vandepitte, D.
van de Par, Steven
Van Dijk, Pim
van Lieshout, P.
van Santen, Jan P.
van Schijndel, Nicolle
van Son, R.
Van Tasell, Dianne
Van Veen, Barry
Vatikiotis-Bateson, Eric A.
Vercoe, Berry
Verdier, Claude
Verge, Marc-Pierre
Verrillo, Ronald T.
Vidmar, Paul
Vieira, M.
Viemeister, Neal F.
Vikman, E.
Vilkman, Erkki
Vorlaender, Michael
Voronovich, Alexander G.
Wada, Hiroshi
Wade, Glen
Wadley, Haydn N. G.
Wagner, James W.
Wakefield, Gregory
Wang, Taylor
Wang, Xioqin
Wang, Y. J.
Wang, Lily M.
Ward, Darren B.
Ward, William
Watkins, William
Watson, Charles S.
Waxler, Roger
Weaver, Richard L.
Weglein, Arthur B.
Weinreich, Gabriel
Weisenberger, Jan M.
Weismer, Gary G.
Weiss, Anthony J.
Wenzel, Alan R.
Wenzel, Elizabeth M.
Werner, Lynne A.
Westbury, J.
Westwood, Evan
Whalen, D.
White, Michael J.
Whitehead, Harold
Whitworth, Glen
Wightman, Frederic L.
Wilen, Larry A.
Wilhelms-Tricarico, Reiner
Williams, David
Williams, Earl G.
Williams, Kevin
Wilson, D. K.
Wilson, O. B., Jr.
Winsberg, Suzanne
Wiskin, James
Witten, Allan J.
Wogram, Klaus
Wolf, Stephen
Wong, G. S. K.
Wong, Bor-Tsuen
Woodhouse, James
Worcester, Peter F.
Wouters, Jan
Wright, Beverly A.
Wright, Frederic L.
Wright, Howard
Wright, Wayne
Wu, Chao-Min
Wu, Cheng Yi
Wu, Enboa
Wu, Junru
Wu, Sean F.
Wu, Ting-Wen
Wurmser, Dan
Wyngaard, John
Xu, Pei-cheng
Xu, Yi
Yairi, E.
Yama, Mark
Yamamoto, Tokuo
Yang, Tsih C.
Yano, Hiroyuki
Yasuda, Kenji
Yates, Graeme
Yehia, Hani
Yoon, Suk Wang
Yoshikawa, Shigeru
Yoshioka, H.
Yost, William A.
Yu, S. S.
Yu. Yung H.
Yun, Kwang S.
Zabolotskaya, Zhenia
Zahorian, Stephen A.
Zahorik, Pavel
Zajac, David J.
Zakarauskas, Pierre
Zattore, Robert
Zeng, Fan-Gang
Zhang, Bixing
Zuckerwar, Allan J.
Zurek, Patrick
Zwolan, Terry

Regional Chapter News

Madras, India: At its meeting on 8 March 2000, Madras-India Regional Chapter of the Acoustical Society of America (MIRC-ASA) presented an Honorary Membership Certificate of the MIRC-ASA and the Mira Paul memorial award of the Acoustical Foundation to Professor Monohar L. Munjal (see Figs. 1 and 2). Professor Munjal is President of the Acoustical Society of India.

FIG. 1. Trustees of the MIRC-ASA present Honorary Membership Certificate to Professor M. L. Munjal. Left to right: P. S. Muthukumar, A. Ramachandraiah (Treasurer), H. S. Paul (Secretary and Chapter Representative), M. L. Munjal, C. P. Vendhan (President), and S. Swarnamani (Vice-President).

FIG. 2. MIRC-ASA Treasurer Professor A. Ramachandraiha (r) presents Mira Paul memorial award of the Acoustical Foundation to Professor M. L. Munjal (l).

Errata

Two errors were made in the caption for the photograph of new Fellows of the Society [President's Report on the 138th Meeting, J. Acoust. Soc. Am. **107**, 2314 (2000)] that appeared as Fig. 5 in the Acoustical News U.S.A. section. In the back row, the person identified as Richard H. Talaske is actually R. Lawrence Kirkegaard who received his ASA Fellowship certificate at the Fall 1999 meeting of the Society. Also, the middle initial for James W. Cottingham was incorrectly listed as ''M.'' The editor apologizes to these individuals for the errors.

USA Meetings Calendar

Listed below is a summary of meetings related to acoustics to be held in the U.S. in the near future. The month/year notation refers to the issue in which a complete meeting announcement appeared.

2000

4–8 December	Joint Meeting: 140th Meeting of the Acoustical Society of America/NoiseCon 2000, Newport Beach, CA [Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; WWW:asa.aip.org].

2001

4–8 February	Midwinter Meeting, Association for Research in Otolaryngology, St. Petersburg, FL [ARO Office, 19 Mantua Rd., Mt. Royal, NJ 08061; Tel: 856-423-7222; Fax: 856-423-3420; E-mail: meetings@aro.org; WWW: www.aro.org/mwm/mwm.html].
22–25 March	''New Frontiers in the Amelioration of Hearing Loss,'' St. Louis, MO [Sarah Uffman, CID Department of Research, 4560 Clayton Ave., St. Louis, MO 63110; Tel.: 314-977-0278; Fax: 314-977-0030; E-mail: suffman@cid.wustl.edu].
30 April–3 May	2001 SAE Noise & Vibration Conference & Exposition, Traverse City, MI [Patti Kreh, SAE Int'l, 755 W. Big Beaver Rd., Suite 1600, Troy, MI 48084; Tel.: 248-273-2474; Fax: 248-273-2494; E-mail: pkreh@sae.org].
4–8 June	141st Meeting of the Acoustical Society of America, Chicago, IL [Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; WWW: asa.aip.org].
15–19 August	ClarinetFest 2001, New Orleans, LA [Dr. Keith Koons, ICA Research Presentation Committee Chair, Music Dept., Univ. of Central Florida, P. O. Box 161354, Orlando, FL 32816-1354; Tel.: 407-823-5116; E-mail: kkoons@pegasus.cc.ucf.edu]. Deadline for receipt of abstracts: 15 January 2001.
7–10 October	2001 IEEE International Ultrasonics Symposium Joint with World Congress on Ultrasonics, Atlanta, GA [W. O'Brien, Electrical and Computer Engineering, Univ. of Illinois, 405 N. Mathews, Urbana, IL 61801; Fax: 217-244-0105; WWW: www.ieee-uffc.org/2001].
3–7 December	142nd Meeting of the Acoustical Society of America, Ft. Lauderdale, FL [Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; WWW: asa.aip.org].

Revision List

New Associates

Acker, Barbara E., 247 Green St., Auburn, AL 36830
Agness, John R., 24872 Sausalito St., Laguna Hills, CA 92653
Akamatsu, Tomonari, National Research Inst. of Fisheries, Engineering, Ebidai Hasaki, Kashima, Ibaraki 314-0421, Japan
Allan, Grant, 10 Taupo Tce., Feilding, New Zealand
Anstee, Stuart D., 39 Polding St., Fairfield Heights, NSW 2165, Australia
Beach, Kirstin E., Dept. of Communicative Disorders, Univ. of Wisconsin—Stevens Point, 1901 4th St., Stevens Point, WI 54481
Becker, Gunnar R., Erlenstrasse 127A, Bremen 28199, Germany
Bednarik, Michael, Dept. of Physics, Czech Tech. Univ. in Prague, Technicka 2, Prague 16627, Czech Republic
Benson, Jonathan W., SAIC/Demaco, 100 Trade Center Dr., Ste. 303, Champaign, IL 61820
Blackwell, Susanna B., 120 Tamarack Dr., Aptos, GA 95003
Boler, Eric, SAIC, 474 Hewes Ave., Ste. 105, Gulfport, MS 39507
Brandmaier, Karl W., AudioMed, Inc., 16043 Valley View Ave., Sante Fe Springs, CA 90670
Brandstein, Michael S., Div. of Engineering and Applied Science, Harvard Univ., 33 Oxford St., Cambridge, MA 02138
Burnett, Gregory C., 675 S. ''H'' St., Livermore, CA 94550
Byrd, Matthew C., Headsets, Inc., 2320 Lakeview Dr., Amarillo, TX 79109
Carpenter, Dave, 5030 N. Nashville Ave., Chicago, IL 60656
Cartes, David A., 222a Campbell St., White River Junction, VT 05001
Chan, Desmond D. B., Acoustics and Air Testing Lab. Company Ltd., Unit 1501A, 15/F Landmark N., Sheung Shui, N.T. Hong Kong
Choe, Howard C., 8 Tanglewood Way N., Andover, MA 01810
Chong, Fan, Acoustics and Air Testing Lab. Company Ltd., Unit 1501A, 15/F Landmark N., Sheung Shui, N.T. Hong Kong
Christine, John T., InServ, General Construction, 121 Dickens Rd., Fuquay-Varina, NC 27526
Corey, John A., CFIC, Inc., 302 Tenth St., Troy, NY 12180
Covey, Joel F., Oceanic & Naval Systems, Northrup Grumman Corporation, P. O. Box 1488, MS 9845, Annapolis, MD 21404
Daniel, Arthur L., Compaq Computer Corp., 20555 SH 249, Houston, TX 77070-2698
Day, Joseph L., 915 Sloan Ave., Waldorf, MD 20602
De Kleine, Emile, Otorhinolaryngology Dept., Univ. Hospital Groningen, Hanzeplein 1, Groningen 9713GZ, The Netherlands
DeBiasi, Marco T., 4206 Palo Verde Rd., Irvine, CA 92612
Dennis, John A., Radiance Technologies, Inc., 500 Wynn Dr., Ste. 504, Huntsville, AL 35816
Eaton, Stuart H., Engineering Section, Prevention Div., Workers' Compensation Board of British Columbia, 8100 Granville Ave., Richmond, BC V6Y 3T6, Canada
Ekhaus, Ira B., Synap Corporation, 100 Wildwood Ave., Arlington, MA 02476
Erwin, Harry R., 13121 Applegrove Lane, Herndon, VA 20171
Friesel, Mark A., 217 Pingree Ave., Ewing, NJ 08618
Garcia, Jose M., Avda. de los Caidos, 4 7 izqda., 15009 A Coruna, Spain
Garlick, Arnold L., Pacific Search Consultants, 2377 S. El Camino Real, #201, San Clemente, CA 92672
Geiger, William H., WHG & Associates, 2421 Edgehill Rd., Toledo, OH 43615
Ghiotto, Alessandro, Nautronix Ltd., 108 Marine Tce., Fremantle WA 6160, Australia
Gordon, Matthew K., Linguistics Dept., Univ. of California, Santa Barbara, Santa Barbara, CA 93106
Gori, Paola, Electronics, Univ. of Roma Tre, Via Della Vasca Navale 84, Rome 00186, Italy
Gorling, R. L. A., RLA Gorling, RR 1, 18236 Glen Rd., Williamstown, ON K0C 2J0, Canada
Grove Deborah M., Microphone Array Group, CAIP, CORE Bldg., 96 Frelinghuysen Rd., Rutgers Univ., Piscataway, NJ 08854-8088
Ha, Sung K., Mechanical Engineering, Hanyang Univ., Sai-dong, Kyungki-do, Ansan 425-791, Korea
Hackman, David G., 11 Gladstone Rd., Marlton, NJ 08053
Hamilton, Miriam, Science Applications Intl. Corp., Advances Systems Group, 4001 Fairfax Dr., Ste. 600, Arlington, VA 22203
Harley, Heidi E., New College of U.S.F., Div. of Social Sciences, 5700 N. Tamiami Trail, Sarasota, FL 34243
Hashim, Zulkifli B., Juru & Teras Bersekutu Sdn Bhd, 3-1A Jalan Pandan Indah 4/1, Taman Pandan Indah, Kuala Lumpur, Wilayah Persekutuan, 55100 Malaysia
Hayden, Charles S., NIOSH, DART/EHPB/HLPS, 4676 Columbia Pkwy., C27, Cincinnati, OH 45226

Hedqvist, Andreas E., TOS-OF, European Space Operations Centre, Robert Bosch Strasse 5, Darmstadt 64293, Germany
Hiroyuki, Okubo, NHK Science and Technical Research Labs., 3D Audio Visual Systems Res. Div., 1-10-11 Kinuta, Setagaya, Tokyo 157-8510, Japan
Hosseini, Seyed Hamid Reza, Shock Wave Research Center, Inst. of Fluid Science, Tohoku Univ., 2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan
Hughes, Michael S., 1632 Ridge Bend Dr., Glencoe, MO 63038
Hulce, Emily A., Arup Acoustics, The Tower, Melbourne Central, Floor 12, 360 Elizabeth St., Melbourne VIC 3000, Australia
Hustvedt, David, AcoustiFLO, 5541 Central Ave., Ste. 125, Boulder, CO 80302
Hwang, Ho-Jin, 7272 E. 37th St. N., #622, Wichita, KS 67226
Ichimura, Hideyuki, 6-49-17-210 Higashi-Nippori, Arakawa-ku, Tokyo 116-0014, Japan
James, Adrian W., 65 Yarmouth Rd., Blofield, Norwich, Norfolk NR13 4LG, U.K.
Jensen, Eric R., 380 Willow Rd., Menlo Park, CA 94025-2714
Jeon, Jin Yong, School of Architectural Engineering, Hanyang Univ., 17 Haengdangdong, Sung Dong-Ku, Seoul 133-791, Korea
Kadman, Yoram, 3 Stern St., Herzlia, Neve Amirim 46412, Israel
Kane, M. R. Schlumberger-Doll Research, Old Quarry Rd., Ridgefield, CT 06877
Karklins, Janis 2671 Lasalle Way, San Jose, CA 95130
Kawamura, Satoru, Shimochujocho 4-13, Ibaraki, Osaka 567-0886, Japan
Kello, Christopher T., House Ear Inst., 2100 W. Third St., Los Angeles, CA 90057-1922
Kelly, Ellen, 11 Kensington Court, Asbury, NJ 08802
Kok, Ben H. M., PBRI/SIAP, P. O. Box 720, Uden 5400-AS, The Netherlands
Koptenko, Sergei V., 189 Felix Rd., Richmond Hill, ON L4C 3M8, Canada
Korporaal, Matthew T., Orincon Corporation, Software Engineering, 9363 Towne Centre Dr., San Diego, CA 92121
Larson, Martha M., 1538 W. Henderson St., Chicago, IL 60657
Larson, Vernon D., Audiology & Acoustics Lab., Howard Leight Industries, 7828 Waterville Rd., San Diego, CA 92173
Latzel, Matthias, Interton Hoergeraete GmbH, Audiology Dept., Am Daennekamp 15, Bergisch Gladbach D-51469, Germany
Lee, Jaiyong, 12467 Barry Knoll Ln., Houston, TX 77024
Leung, Yuen Tat Y. T., Acoustics and Air Testing Lab. Company Ltd., Unit 1501A, 15/F Landmark N., Sheung Shui, N. T. Hong Kong
Lydolf, Morten, Dept. of Acoustics, Aalborg Univ., Fredrik Bajers Vej 7-B4, Aalborg DK-9220, Denmark
Ma, Li, Inst. of Acoustics, Academia Sinica of China, 17 Zhongguancun St., Beijing 100080, China
Maltsev, Nick E., 10678D Maplewood Rd., Cupertino, CA 95014
Manneville, Sebastien, Lab. Ondes et Acoustique, Ecole Superieure de Physique et, Chimie Industrielles, 10 rue Vauquelin, Paris 75005, France
Maralina, Galina V., P. O. Box 500, Oregon House, CA 95962
Marulo, Francesco, Via vecchia Aquini 8, Boscoreale, Naples 80041, Italy
Mason, Andrew C., Div. of Life Sciences, Univ. of Toronto at Scarborough, 1265 Military Trail, Scarborough, ON M1C 1A4, Canada
Mattys, Sven L., House Ear Inst., 2100 W. Third St., Los Angeles, CA 90057
Mauroy, Richard E., Rue de Monds 24, Braine Le Comte B-7090, Belgium
McHenry, Monica A., Communication Disorders, Univ. of Houston, 4800 Calhoun Blvd., Houston, TX 77204-6611
Mease, Michael R., 707 Pittman Dr., Sumter, SC 29154
Mehta, Sandip R., Collins & Aikman, 47785 W. Anchor Court, Plymouth, MI 48170
Metcalf, Randy B., P. O. Box 7075, Pensacola, FL 32534
Metzger, Nicole S., Tyco Submarine Systems, Ltd., Engineering and Construction, Cable Route Survey, 60 Columbia Turnpike, Bldg. A, Room 2A-145, Morristown, NJ 07960
Michiko, Nagayoshi, 2-2-1-1116 Motomachi, Hakata-ku, Fukuoka City 816-0077, Japan
Moreno, Jaime, 15215 Victory Blvd., #109, Van Nuys, CA 91411
Moy, Jorge C., Av. Ameghino, K-5 Urb. La Calera de la Merced, Lima, Peru
Nakai, Takayoshi, Dept. of Electrical and Electronic Eng., Shizuoka Univ., 3-5-1 Johoku, Hamamatsu, Shizuoka 432-8561, Japan
Norwood, Ronald F., 1800 Silent Forest Point, Oviedo, FL 32765
O'Brien, Michael W., Electrical and Electronic Eng., Cork Inst. of Technology, Rossa Ave., Bishopstown, Cork, Munster, Ireland
Oh, Sei-Jin, Acoustics Electronics Eng., JuSeong College, Mt. 4 Duckman-Ri NaeSu-Ub, Cheong Wong Gun, Chung Buk, 363-794, South Korea
Ojanen, Ossi T., Ketostenkatu, Tampere FIN-3300, Finland
Pacull, Michel, Aerospatiale Matra Airbus, Acoustics Dept., 316 Route de Bayonne, BP M0142/4, Toulouse F-31060, Cedex 06, France
Pai, Dinesh K., Computer Science, Univ. of British Columbia, 2366 Main Mall, Vancouver, BC V6T 1Z4, Canada
Peiffer, Alexander, Rolandstr. 104, Berlin 13156, Germany
Pettersen, Michael S., 2737 Garrison Ave., Evanston, IL 60201
Ponirakis, Dimitri W., 2715 Jones St., Sioux City, IA 51104
Rao, Ashwin P., 640 Worcester Rd., Framingham, MA 01702
Rigaudias, Thibault, SAGEM, Queen Charlotte House, Chatman Maritime, Kent NE4 4QZ, U.K.
Ross, David, P. O. Box 544, Rumson, NJ 07760
Sabrin, Howard W., 624 Concerto Ln., Silver Spring, MD 20901
Sakagami, Kimihiro, Environmental Acoustics Lab., Faculty of Engineering, Kobe Univ., 1-1 Rokkodai, Nada, Kobe, Hyogo 658-0022, Japan
Sakakibara, Ken-Ichi, NTT Communication Science Labs., 3-1 Morinosato Wakamiya, Atsugi-shi, Kanagawa 243-0198, Japan
Salameh, Mohamad, 712 E. St. Andrews, Midland, MI 48642
Sauer, Rick A., 130 Camp Ground Rd., Brownsboro, AL 35741
Shaffer, Don L., 461 Mercer St. NW, Palm Bay, FL 32907
Simon, Jonathan Z., Inst. for Systems Research, Univ. of Maryland, College Park, MD 20742
Solokhin, Nickolay V., 1174 Walker Ave., #3, Walnut Creek, CA 94596
Soto Nicolas, Alberto, Nokia Mobile Phones, P. O. Box 83, Tampere 33721, Finland
Stocker, Michael, P. O. Box 559, Lagunitas, CA 94938
Strange, Shawn, Aetc, Inc., 8910 University Center Ln., Ste. 900, San Diego, CA 92122
Sutin, Alexander M., Davidson Lab., Stevens Inst. of Technology, 711 Hudson St., Hoboken, NJ 07030
Tabei, Makoto, Electrical and Computer Eng., Univ. of Rochester, 3-4309, MED CTR Box 648, 601 Elmwood Ave., Rochester, NY 14642
Taylor, Paul, Torrence Sound Equipment Company, 29050 Glenwood Rd., Perrysburg, OH 43551
Telling, Rob H., Environmental Acoustics, Newtons Rd., Weymouth, Dorset DT4 8UR, U.K.
Temperli, Philippe, 13 Ch. Union, Savigny VD 1073, Switzerland
Thiel, David D., Microsoft, One Microsoft Way, Redmond, WA 98052
Thomenius, Kai E., 74 Van Vranken Rd., Clifton Park, NY 12065-6804
Thurgood, Elzbieta A., Dept. of English, California State Univ., Chico, CA 95920-0830
Tran, Joseph B., Acoustic Lab., Soundcoat, 16901 Armstrong Ave., Irvine, CA 92606
Tristano, Giovanni, Via Lazio 92, Palermo, Sicily I-90144, Italy
Tryanski, Danielle, Research and Technical Services, Starkey Laboratories, 6600 Washington Ave. S., Eden Prairie, MN 55344
Valle, Christine, Mechanical Engineering, Univ. of Maine, 212 Boardman Hall, Orono, ME 04469-5711
van Dommelen, Wim A., Dept. of Linguistics, Norwegian Univ. of Science and Technology, Trondheim N-7491, Norway
Vincenot, Eric, NEXO, BP 50107, Roissy, CDG, Cedex F-95950, France
Wakefield, Gregory H., Electrical Engineering and Computer Science, Univ. of Michigan, 1301 Beal Ave., Ann Arbor, MI 48109
Waldmann, Bernd, Stolzingstr. 8, Munchen D-81927, Germany
Wang, Chong, Apt. 807, 1525 Alta Vista Dr., Ottawa, ON K1G 0G1, Canada
Waters, Timothy R., McKay Conant Brook, Inc., AV Group, 5655 Lindero Canyon Rd., Ste. 325, West Lake Village, CA 91362
Welti, Todd S., Sonjavej 38D, Aalborg, Jutland 9000, Denmark
Willatzen, Morten, Mads Clausen Inst., Univ. of Southern Denmark, Grundtrigs Alle 150, Sunderbury 6400, Denmark
Wright, Timothy F., Dept. of Psychology, Univ. of Maryland, College Park, MD 20742
Wu, Weixiong, Allee King Rosen & Fleming, Inc., 117 E. 29th St., New York, NY 10016
Young Heon, Kwon, Hai Chung Co., Ltd., C.P.O. Box 7329, Seoul 100-673, South Korea

Younger, Michael B., BAE Systems, Inc., 6500 Tracor Ln., 27-17, Austin, TX 78725
Yukio, Iwaya, Dept. of Electrical and Electronic Eng., Akita Univ., 1-1 Tegata Gakuencho, Akita 010-8502, Japan
Zeegers, Jos, Eindhoven Univ. of Technology, Physics Dept., Den Dolech 2, P. O. Box 513, Eindhoven NL-5600MB, The Netherlands
Zurk, Lisa M., Advanced Array Technology, MIT Lincoln Lab., 244 Wood St., Lexington, MA 02420

New Students

Abrams, Daniel A., Northwestern Univ., Hearing Science and Audiology, 2299 North Campus Dr., Evanston, IL 60208
Adam, Trudy, J., Audiology and Speech Sciences, Univ. of British Columbia, 226-5804 Fairview Ave., Vancouver, BC V6T 1Z3, Canada
Agarwal, Meena, Dept. of Speech Language Pathology, Bowling Green State Univ., Bowling Green, OH 43403
Alexander, Joshua M., 7085 Watts Rd., Madison, WI 53719
Arifianto, Dhany, Tokyo Inst. of Technology, Information Processing Lab., 4259 Nagatsuta-cho, Midori-ku, Yokohama-shi, Kanagawa-ken 2260-8502, Japan
Avelino, Heriberto, Linguistics Dept., UCLA, 3125 Campbell Hall, Hilgard Ave., Los Angeles, CA 90095
Azrael, 6608 Strutton Cove, Austin, TX 78759
Balachandran, Rupa, 34 Anthony Ln., Lawrenceville, NJ 08648
Bartsch, Guido, Technical Univ. of Aachen, Inst. of Technical Acoustics, Templergraben 55, Aachen 52056, Germany
Bauer, Jay J., Communication, Sciences & Disorders, Northwestern Univ., 2299 North Campus Dr., Evanston, IL 60208
Bentrem, Frank W., 6490 US Highway 49, Apt. 291, Hattiesburg, MS 39401-3013
Blanc, Arthur, Mechnical Engineering, Virginia Tech, ME VAL NEB 142, Blacksburg, VA 24061
Bolton, Dean, 64 Bay State Rd., Boston, MA 02215
Cartes, David A., 222A Campbell St., White River Junction, VT 05001
Cavagnaro, Laura M., Center for Coastal and Ocean Mapping, Univ. of New Hampshire, 24 Colovos Rd., Durham, NH 03857
Chen, Wen-Shiang, Bioengineering Dept., Applied Physics Lab., 1013 NE 40th St., Seattle, WA 98105
Cichock, Joseph A., 2311 NE 4th Ave., Boca Raton, FL 33431
Connaghan, Kathryn P., 5319 Keystone Place N., Seattle, WA 98103
Curtin, Suzanne L., Linguistics Dept., Univ. of Southern California, GFS 301, University Park, Los Angeles, CA 90089-1963
Davis, Carrie J., Speech and Hearing Sciences, Indiana Univ., 200 S. Jordan, Bloomington, IN 47405
Diakoumakou, Evanthia, Linguistics Dept., Univ. of Michigan, 1076 Frieze Bldg., Ann Arbor, MI 48109-1285
Dietz, Jerrold E. G., Dept. of Ocean Engineering, Seatech c/o Florida Atlantic Univ., 101 N. Beach Rd., Danie Beach, FL 33004
Dodge, Leanne E., 3005 Winston St., Bloomington, IN 47401
Ellis, Donna A., 521 Manor Rd., Severna Park, MD 21146
Fan, Yun-Hui, 3450 Evans Rd., Apt. 105D, Chamblee, GA 30341-5823
Fedewa, Russell, J., Washington Univ., Physics Box 1105, One Brookings Dr., St. Louis, MO 63130
Fei, Dong, Center for Nondestructive Evaluation, Iowa State Univ., 137 ASCII, 1915 Scholl Rd., Ames, IA 50011
Gant, Valdez L., 611 Risen Star Dr., Crestview, FL 32539
Gattemy, Glen L., 708 Caine, Socorro, NM 87801
Goodman, Shawn S., 601 Redbud Hill, Bloomington, IN 47408
Harrison, Byron W., 95 Evergreen Ave., Hartford, CT 06105
Helser, Janelle L., 256 Corl St., #2B, State College, PA 16801
Hernandez Morales, Fermin H., c/o Corregidor Jose de Pasamonte No. 12, Madrid 28030, Spain
Hill, Rebecca B., 8 SW 10th St., Gainesville, FL 32601
Hirst, Jonathan M., 40 St. Werbrughs Rd., Manchester M21 0TJ, U.K.
Hoelting, Marcia K., 1005 Elmwood Cove, Cedar Park, TX 78613
Hojen, Anders, English Dept., Univ. of Aarhus, Jens Chr. Skous Vej 7, Aarhus 8000, Denmark
Hu, Changmin, School of Mechanical and Production Eng., Nanyang Technological Univ., Nanyang Ave., Mechanics of Machines Lab., Singapore 639798
Hwang, George, 1321 Donal Dr., Flint, MI 48532
Kelly, Kate, 1001 W. Mill St., Heber Springs, AR 72543
Keranen, Joe G., 2604 Plaza Dr., State College, PA 16801
Kim, Sang Kun, 94 Abercrombie St., Chipperdale, Sydney NSW 2008, Australia
Kinney, Ashlynn L., Dept. of Psychology, Univ. of Texas at Austin, Austin, TX 78712
Kollman, Steven A., Electronic Technology, Gateway Technological College, 1001 Main St., Racine, WI 53403
Kourakata, Itaru, 9343-10 Ikarashi, 3-nochou, Nigata City 950-2171, Japan
Kovacic, Damir, Intl. School for Advanced Studies/SISSA, Neuroscience Sector, via Beirut 2-4, Trieste 34014, Italy
Kroeff, Gia, Aeronautical and Automotive Dept., Loughborough Univ., Leicestershire LE11 3TU, U.K.
Laabs, Jeremy K., 8193 Via Mallorca, La Jolla, CA 92037
Lai, Yi-San, 950 Massachusetts Ave., Apt. 515, Cambridge, MA 02139
Lane, Courtney C., Massachusetts Eye and Ear Infirmary, Eaton Peabody Lab., 243 Charles St., Boston, MA 02114
Lee, Chao-Yang, Cognitive and Linguistic Sciences, Brown Univ., Box 1978, Providence, RI 02912
Lee, Young-Sup, ISVR, Univ. of Southampton, Highfield, Southampton, Hampshire SO17 1BJ, U.K.
Lichtenhan, Jeffrey T., 722 Lincoln, Wamego, KS 66547
Liljegren, Kristin, 4304-D SW 70th Terrace, Gainesville, FL 32608
Loebach, Jeremy L., 812 E. Kerr Ave., #206, Urbana, IL 61802
LoRang, John M., 835 NE 3rd Ave., Gainesville, FL 32601
Martin, Geoff G., Multichannel Audio Research Lab., McGill Univ., 555 Sherbrooke St. W., Montreal, QC H3A 1E3, Canada
Moffitt, Jeffrey R., Physics Dept., The College of Wooster, 1189 Beall Ave., Wooster, OH 44691
Mondini, Michele, Dept. of Psychology, Northeastern Univ., 125 Nightingale Hall, Boston, MA 02115
Mueller, Tobias M., Zaehringerstr. 13, Berlin 10707, Germany
Najolia, Dorie A., 415 S. Atherton St., C-6, State College, PA 16801
Napoletano, Brian M., Dept. of Entomology, Computational Ecology and Visualization Lab., 235-B Natural Sciences Bldg., Michigan State Univ., East Lansing, MI 48824
Nardi, Flavio, Georgia Inst. of Technology, Aerospace Engineering, Montgomery Knight Bldg., Atlanta, GA 30332
Nelson, Patrick J., 399 Main St., Saco, ME 04072
Newton, Jessica E., Box 1983, 200 Bloomfield Ave., West Hartford, CT 06117
Nissen, Shawn, 1129 Selis Ave., Apt. C, Columbus, OH 43212
Noad, Michael J., 36 Fairscence Crescent, Avoca NSW 2251, Australia
Oishi, Tetsuro, ECE Dept., Univ. of Massachusetts, Dartmouth, 285 Old Westport Rd., North Dartmouth, MA 02747
Oleson, Erin M., Scripps Inst. of Oceanography, Univ. of California, San Diego, 9500 Gilman Dr., #0208, La Jolla, CA 92093
Olson, Nicole G., 115 Oakland St., Berkeley Heights, NJ 07922
Othman, Shadi F., IIT, Mechanical Materials & Aerospace, E1 Bldg., 10 W 32nd St., Chicago, IL 60616
Painter, Ted, 106 W. Acton Rd., Stow, MA 01775
Pardo, Bryan A., Artificial Intelligence Lab., EECS, Univ. of Michigan, 110 ATL Blvd., 1101 Beal Ave., Ann Arbor, MI 48109-2110
Park, Hansang, 3352 Lake Austin Blvd., Apt. C, Austin, TX 78703
Park, Junhong, Mechanical Engineering, Purdue Univ., 1077 Ray W. Herrick Lab., West Lafayette, IN 47907-1077
Park, Sewon, Pennsylvania State Univ., Applied Research Lab., Atherton St., State College, PA 16801
Perez, Hernan E., Lab. de Fonetics, Univ. de Concepcion, Casilla 120, Correo 3, Concepcion, Chile
Pilkington, Wayne C., 165 Trevor Court Rd., Rochester, NY 14610
Pompei, F. Joseph, Media Arts and Sciences, Massachusetts Inst. of Technology, 20 Ames St., E15-401, Cambridge, MA 02139
Ratilal, Purnima, 143 Albany St., #134A, Cambridge, MA 02139-4262
Reed, Allen H., Naval Research Lab., Marine Geosciences, Bldg. 1005, Code 7431, Stennis Space Center, MS 39529
Richey, Colleen, Linguistics Dept., Stanford Univ., Margaret Jacks Hall, Bldg. 460, Stanford, CA 94305-5666
Roginska, Agnieszka, Music Technology, Northwestern Univ., 711 Elgin Rd., Evanston, IL 60208
Rollo, Mario M., Museu de Zoologia-USP, Mastozoologia, Av. Nazare 481, Sao Paulo 04263-000, Brazil

Rotter, Sarah M., 11917 Villa Dorado Dr., St. Louis, MO 63146
Sabol, Bruce M., 221 Funcher Rd., Vicksburg, MS 39180
Sanchez, Robert J., 1410 Palo Verde Dr., Irvine, CA 92621
Sarampalis, Anastasios, 69 Wimpole Rd., Colchester C01 2DW, U.K.
Seablom, Danielle M., Community Baptist Church, CBC Tech, 9090 19th St., Alta Loma, CA 91701
Shi, Yijun, Nondestructive Evaluation Lab., Massachusetts Inst. of Technology, Room 1-065, Cambridge, MA 02139
Sinha, Naveen N., 112 Shirlane Place, Los Alamos, NM 87544
Sousa, Ana VG, Rua Ney Armando Meziat, 31/105-Ilha do Governado, Rio de Janeiro 21941-090, Brazil
Spiegelberg, Scott, 283 Main St., Tonawanda, NY 14150
Stainsby, Thomas H., 1371 Dandenong Rd., Armadale VIC 3143, Australia
Stefaniw, Abigail E., 2165 14th St., Troy, NY 12180
Stickney, Ginger S., Human Development and Communication Sciences, Univ. of Texas at Dallas, GR41, Box 830688, Richardson, TX 75083-0688
Streeter, Timothy M., Cognitive and Neural Systems, Boston Univ., 677 Beacon St., Boston, MA 02215
Su-Hyun, Jin, 1405 Jessamine Ave. W., #208, St. Paul, MN 55108
Sun, Xuejing, Communication Sciences and Disorders, Northwestern Univ., 2299 North Campus Dr., Evanston, IL 60208-3570
Tachner, Joyce L., Acoustics, 17961 Sky Park Circle, Irvine, CA 92614
Tirtabudi, Prisilia, Psychology Dept., California State Univ. Los Angeles, 5151 State University Dr., Los Angeles, CA 90032
Trousil, Rebecca L., Physics Dept., Washington Univ., One Brookings Dr., Campus Box 1105, St. Louis, MO 63130
Tsakostas, Christos, Digeni Akrita 29, Athens 12243, Greece
Tserdanelis, Georgios, Ohio State Univ., Dept. of Linguistics, 1712 Neil Ave., Columbus, OH 43210
Vester, Heike I., Dept. of Arctic Biology, Univ. of Tromsoe, Breivika, Tromsoe, Troms 9037, Norway
Volaric, Steve J., 1260 N. Bosworth Ave., Chicago, IL 60622
Weiland, Nathan, Mechanical Engineering, Georgia Tech, 333600 Georgia Tech Station, Atlanta, GA 30332-1505
Winker, Douglas F., Electrical and Computer Engineering, Univ. of Texas at Austin, 12007 Rotherham, Austin, TX 78753
Wunderlich, Adam J., Electrical and Computer Engineering, Bioacoustics Research Laboratory, B428C Beckman Inst., MC 251, 405 N. Mathews, Urbana, IL 61801
Xu, Wen, 540 Memorial Dr., Apt. 408, Cambridge, MA 02139
Yan, Zhongyu, 505 Sandheger Place, Apt. 5, Cincinnati, OH 45220-2744
Yuan, Jiahong, Dept. of Linguistics, Cornell Univ., Morrill Hall, Ithaca, NY 14853

Students to Associates

C. M. Elliott, K. H. Franck, A. Ghandi, R. Guedes, R. M. Hemphill, F. Iglehart, J. C. Jankovsky, P. S. Deliwala, E. A. Macpherson, N. Smith, R. G. Tull, J. T. Viechnicki

Associates Elected Members

N. Aizawa, M. Ansorge, R. J. Barton, M. Bednarik, R. S. Bolia, M. W. Bonilha, S. D. Brady, L. Brancazio, M. S. Brandstein, G. Chen, R. O. Cleveland, A. P. Croney, D. T. DiPerna, W. R. Drennan, F. A. Dube, H. W. Gierlich, G. M. Glickman, S.-K. Han, N. B. Higbie, L. Hoff, J. M. Impagliazzo, A. W. James, S. W. Kahn, J. Kang, E. A. Karlow, N. J. Kessissoglou, J. A. Ketterling, V. D. Larson, A. C. Lavery, J. Lee, V. V. Lenchine, J. J. Lister, T. M. Logan, G. J. Macaulay, R. Makarewicz, S. Mazzola, D. W. Merdes, M. W. McNeary, S.-J. Moon, S. J. Page, S. Petrovic, E. S. Prame, J. L. Punch, R. D. Read, E. L. Reuter, J. S. Ryder, C. Shera, I. Shiro, P. J. Shorter, W. Sun, J. A. Thomas, S. Tokuma, L. J. Trainor, J. R. Walker, G. Whitworth, S. T. Roweis, L. H. Sibul, H. C. Strifors, T. Wang, J. Weinstein, P. Zhang

Members Elected Fellows

R. D. Ayers, S. Bech, D. E. Bray, P. D'Antonio, L. J. D. Demany, M. A. Holden, D. E. Hall, G. R. Hamilton, W. J. Hughes, D. F. Jones, R. H. Love, C. Thompson, J-Q. Wang, O. A. Godin, H. S. Paul

Reinstated

M. F. Werby—*Fellow*

T. S. Bragg, S. M. Haas, J. B. Kedzierski, P. Y. Lee, D. Thurmond, J. J. Welsh, H. J. Whitehouse—*Members*

L. R. Brown, A. K. Rogers, R. D. Shilling—*Associates*

Resigned

V. K. Bhat, A. W. D. Jongens, S. Naidoo, J. W. E. Pettersen, L. Seltzer—*Members*

R. L. C. Briers, A. Galembo, B. T. Vo—*Associates*

R. Weber—*Student*

Deceased

D. G. Crighton, J. H. Dewson III, W. F. Meeker, T. W. Tillman, F. E. White—*Fellows*

M. E. Frazer, J. B. Large, H. J. Oyer, R. J. Schwartz, J. H. Thompson—*Members*

F. J. Miller—*Associate*

Dropped

D. A. Berkley, A. Cohen, M. J. Crocker, G. L. Fuchs, T. G. Horwath, G. J. Quentin, M. C. Teich, E. H. Trinch—*Fellows*

O. V. Abromov, Vladimir Abramov, M. C. Allie, Bjorn A. J. Angelsen, M. B. Babbitt, C. D. Barber, T. S. Bell, M. V. Bernblit, S. S. Boatright-Horowitz, J. X. Borzym, John Brady III, D. L. Brewer, Melissa Burn, D. J. Burr, Bio Cai, H. D. Capdevilla, S. J. Caprette, Jr., S. A. Chapman, Francis H. K. Chen, A. N. Chesnokov, J. S. Cornette, L. S. Couchman, J. F. Craynock, L. S. Deliman, S. F. Duncan, W. W. Fearon, W. M. Fisher, R. G. Fizell, J. K. Gaevert, J. N. Gardi, A. N. Gavrilov, R. H. Gikey, Constaintine Glaretas, W. A. Goodwin, F. A. Groenewold, R. H. Gutierrez, K. R. Henry, J. J. Hirschfelder, E. V. Hofbeck, Frederika Holmes, D. E. Homa, Hong Hsu, K. D. Hsueh, A. W. F. Huggins, S. Y. M. Hui, G. D. Hutton, M. D. Johns, M. A. Josserand, L. J. Kelly, Elizabeth Kelly-Fry, S. L. Kirk, Andras Kornai, B. A. Kugler, P. C. C. Lai, R. A. Lapetina, Andes Larraza, Jungmee Lee, Sheau-Fang Lei, Kung-Pu Li, Yuan-Liang Li, Jian-Yu Lin, M. J. Macchi, P. E. Madden, G. G. Martin, K. W. Melville, Kurt Metzger, Jr., M. H. Miller, Stephanie Novak, M. J. O'Kane, M. H. O'Malley, J. E. Oh, M. J. Osberger, L. A. Ostrovskiy, L. E. Pitts, Ulrich Pollmann, Raymond Porzio, John Prohs, M. E. Rothermel, Martti Saldmaa, G. J. Sciacca, Anu Sharma, B. R. Shelton, M. D. Sneddon, C. R. Steele, R. P. Stein, S. A. Stotts, B. M. Sullivan, Nice Terzi, S. K. Tomar, R. J. M. Van Hoelsel, E. J. Voelker, Bary Voroba, R. L. Wisman, E. J. White, W. R. Woszczyk, Weiping Yao, J. W. Young, D. L. Young, Jr., B. L. Zeigler—*Members*

Dino Abate, J. F. Abbott, M. M. Affshar, Pierre Ahad, C. L. Alex, Peer Amble, Richard Auerbach T. C. Austin, Sevig Ayter, G. M. Bachman, Chilki Baek, S. M. Balabaev, P. T. Bangayan, J. P. Barlow, S. M. Barlow, G. E. Bartuska, D. A. Baskiji, L. S. Bernstein, S. E. Blanc, P. A. Blolloch, D. E. Blore, B. J. Bolleman, F. A. Bourge, Willem Brans, D. N. Brass, C. S. Bull, C. R. Burrus, Michelle Caisse, J. R. Callister, N. W. N. Campbell, B. D. Castile, W. A. Castleman, Weicheng Chen, Samior Chennovkh, Wenshyang Chiu, Chang-Hwan Choi, John-Dongwook Chooi, R. W. Cholewiak, R. D. Ciskowksi, Don Colton, L. A. Coury, L. S. Crossland, C. S. Crowther, T. J. Danley, Paolo Dini, Paul Doany, Xingen Dong, J. R. Doucet, J. A. Ekaterinaris, Bartholomew Elias, P. A. Elmore, Alejandro Esbri, Qiping Fang, J. M. Ferrandez, A. A. Ferri, A. A. F. Fichera, L. S. Finegold, K. M. Fire, P. A. Fox, M. S. Fozok, R. K. Furness, P. E. Giva, Jack Green, Marcos Grellet, Boris Gurevich, A. N. Hadjicostis, S. Hahn, K. D. Hall, J. C. T. Hallam, T. A. Hamel, R. B. Hatangadi, M. A. Hawwa, S. P. Hecke, K. E. Hecox, B. S. Hester, D. A. K. Hewlett, Matthias Hey, Farhang Honarvar, L. G. Hopkins, S. A. Hudson, J. B. Ibrahim, S. H. Im, E. M. Jaeger, N. Z. Jamali, R. L. Jenison, Guolinang Jin, C. G. Journeau, R. L. Jueschke, C. W. Kennedy, D. L. Kewley, Tae Kim, B. W. Kimble, Yoichi Kimura, P. J. Klammer, Sverre Knudsen, M. H. Krane, R. S. Kulkarni, Fan-Ming Kuo, D. E. Lambert, G. D. Lame, Edwin Langberg, D. L. Larom, Chang-Myung Lee, J. J. Lee, Ming-Ran Lee, Antonio Leotta, R. A. Lester, H. B. Levene, E. C. Levi,

Pei-Ling Liu, R. A. Mack, D. O. Mackey, A. K. Madan, J. M. Madigan, J. P. Maillard, Marios Mantakas, P. C. Marks, C. D. Martinez, Antonio Mattei, D. W. Mauney, D. J. McDowell, K. G. McGuigan, R. S. McGuinn, E. B. Moltz, D. R. Mook, S. E. Moore, S. M. Morgan, S. I. Muyakshin, Nilabh Narayan, S. P. Neal, K. D. Neltnor, A. J. Napgen, Joel Nguyen, Phuong-Thad N. Nguyen, D. E. Nichols, D. E. Niemoeller, C. E. Nishimura, J. D. Nordman, Eunmi Oh, D. P. Orofino, Djamel Ovis, A. E. Ozelame, Cuneyt Ozturk, Satish Padiyar, J. J. Palakal, G. J. Pan, J. C. Park, P. E. Park, A. D. Patel, Lindamae Peck, Mark Pell, G. K. Penesisi, N. L. Penrose, Dijana Petrovska, D. D. Pierce, J. S. Pruitt, L. O. Ramig, Rama Ratnam, Robertson Rebula, J. D. Reed, T. W. Reeves, R. J. Renomeron, G. H. Robinson, L. D. Robinson, Victoria Rodellar-Biarge, Marc Roelands, N. C. Roussel, C. L. Rowley, H. J. Ruf, Kourosh Saberi, Silvio Santoboni, J. H. Sathoff, S. S. Saunders, August Sauter, Jr., Per Sautermeister, J. C. Scheinber, S. F. Schell, Anthony Schreiner, C. A. Schweitzer, R. C. Shaffer, Edward Shalis, Thom Shaughnessy, J. D. Shelley, R. B. Shuman, L. A. Silvati, M. A. Silviolotti, Harald Slatky, D. J. Smith, J. P. Smith, P. M. Smith, Tao Song, R. J. Soukupu, D. L. Steiger, D. A. Summa, Zhigang Sun, P. H. Szeto, S. M. Tan, S. M. Tarket, Rick Tavares, T. B. Tay, B. J. Taylor, G. R. Terr Haar, C. H. Thompson, D. M. Thompson, G. C. Thompson, G. E. Tisdale, M. W. Todd, Jianming Tsai, Shahram Baezy, A. F. Vakakis, William Vass, N. E. Vaughan, Hernan Vega-Campos, Marc-Pierre Verge, H. T. Vincent, H. I. Vold, J. P. Vuono, Jocelyne Wable, Kuansan Wang, N. J. Wattrus, Reinhard Weber, O. E.Whyche, C. R. Winkel, B. D. Womack, E. W. K. Wong, G. T. Wood, Qiu Wu, S. A. Wymer, B. K. Yeung, Wesley Yi, K. B. Yoo, E. J. Zaiken, Ling Zheng, Xiaoru Zhou, Ruichao Zhu, R. A. Zingarelli, E. C. Zsiga—*Associates*

B. A. Acker, J. N. Anantharaman, Chianing Ao, P. Kaytar, E. W. Biglari, Ch. B. Birdsong, J. R. Blough, Alex Boudreau, M. R. Buehler, E. A. Campos, Jiasu Cao, G. M. Carbone, S. A. Carr, W. B. Shandler, V. V. Cherer-pennikov, Micheal Cheung, Sylvain Clement, A. E. Dainora, Praveen Dalakrishna, Jana Dankovicova, D. H. Dolan, Y. V. Dudko, Jennifer Dugan, M. M. Dupuis, Pierre Elisseeff, C. E. Farell, H. P. Frederickson, Joao Frota, D. E. Gaygen, S. J. Graves, S. J. Guzman, W. H. Ham, C. M. Harne, Daniel Harrison, W. T. Hassan, S. A. Hayes, A. S. Henly, J. K. Henry, S. M. Hirsch, J. E. Horna, C. Q. Howard, D. L. Hughes, S. F. Hynes, Jean-Francois E. Ille, Vijay Jayachandran, L. K. Jin, D. L. Johnson, T. E. Johnston, Jungu Kang, T. D. Kite, Christos Koutsodimakis, J. J. Krogseng, K. C. Kunes, Jaiyong Lee, Samson Lee, M. M. Lentz Chang Liu, P. A. C. Lopes, Albert Lopez-Larrosa, P. A. Machado, J. F. Macneil, N. K. Mandal, E. R. Mansell, Peter Marvit, J. J. Mason, Daiji Mikami, Ramin Mohajer, C. L. Moulton, E. I. Oboznenko, Kaori Omuro, A. G. Petchulscv, Gabriela Petculescu, A. B. Pinus, P. N. Plyler, C. M. Pray, D. H. Reser, R. T. Rhee, Bernd Rieche, Gabriel Ruiz, Peter Rutherford, C. F. Sepulveda, J. B. Shadwick, E. M. Sheppard, C. W. Slater, R. D. Smart, Nadine Smith, Iris Smorodinsky, T. D. Sparks, K. M. Stankovic, E. C. Sullivan, Warren Tam, Robert Tannen, S. A. Van Duyne, G. J. Wadsworth, A. V. Walker, T. N. Wenck, J. C. Williams, A. A. Wilson, Weston Wilson, Ying Xu, Lening Yang, C. L. Yapura, Goangshiuan Ying, Xing Zeng, Jie Zhang—*Students*

Fellows	976
Members	2764
Associates	2793
Students	851
	7384

OBITUARIES

This section of the Journal publishes obituaries concerning the death of Fellows of the Society and other acousticians eminent in the world of acoustics. When notified, the Editor-in-Chief solicits a summary of the person's life and contributions from an ASA member thoroughly familiar with the details, if possible. If a promised obituary is never received, a brief obituary notice may be published later.

Franklin Seaney Cooper • 1908–1999

Franklin S. Cooper, pioneer and educator in the field of speech science, died 20 February 1999 at the age of 90 in Palo Alto, CA. He was a Fellow of the Acoustical Society of America and of the Institute of Electrical and Electronic Engineers (IEEE). He was born 29 April 1908 in Robinson, IL. He would sometimes tell us of his rural upbringing there. His first degree, a B.S. with honors in Engineering Physics, was awarded by the University of Illinois in 1931. His doctorate, a Ph.D. in Physics, was awarded by the Massachusetts Institute of Technology in 1936. His early career was spent as a research engineer at the General Electric Laboratories in Schenectady, NY. In 1935, while still there, he and Caryl P. Haskins founded the Haskins Laboratories as a private not-for-profit research group. That year he married Frances Edith Clem, who died in 1991.

Frank Cooper was a major force in promoting, organizing, and leading early contemporary research into the acoustics of speech production and perception, as well as, later, the study of underlying physiological mechanisms. This happened during the latter part of his stint as Associate Research Director of Haskins Laboratories, 1939–1955, and then as President and Research Director, 1955–1975. In his retirement, he continued to make himself available to the Haskins program as Associate Research Director until 1986, when he and his wife migrated westward. Frank Cooper's other role was that of educator. In the mid-1950s he started opening up the vistas of interdisciplinary speech research to generations of graduate students of linguistics and, later, psychology and speech and hearing.

With limited resources, Frank Cooper's technical originality and ingenuity came to the fore. The invention of the sound spectrograph at Bell Telephone Laboratories, when it became generally available after the Second World War, greatly facilitated acoustic studies of speech. The design was modified by Frank in order to build a photographic spectrograph that produced spectrograms as film transparencies that could be used with his speech synthesizer, the Pattern Playback. He designed the latter to accept either spectrographic patterns by transmission of light through such photographic spectrographs or by reflected light from painted patterns of simplified spectrograms. These patterns could be either derived from real speech or created systematically in steps along selected dimensions of the spectrum to test hypotheses about the information-bearing elements of the speech signal, what we fondly called the acoustic cues. Although other speech synthesizers existed or soon came into being, for example, DAVO of the Massachusetts Institute of Technology, PAT of the Signals Research and Development Establishment and the University of Edinburgh, and OVE II of the Royal Institute of Technology, Stockholm, Cooper's Pattern Playback turned out to be the first one to be used extensively for experiments in speech perception. Frank Cooper also designed other special-purpose synthesizers in the early days: Octopus, Voback, Intonator, and Alexander. Of these four, only the Voback and the Intonator were used extensively for perceptual experiments. They were, in a sense, ''parasitic'' on Homer Dudley's Vocoder. This was all before we had computers. For the execution of these plans, much help was given by the late John M. Borst.

Even in his relentless determination to push the program forward, Frank's benevolent presence and insightful management gave birth to a rich, heady, inspiring atmosphere of interdisciplinary research during the New York City years. Whether working as individuals or as groups of collaborators, the members of the research staff talked to each other quite easily and readily. There was daily interaction and consultation among physicists, engineers, psychologists, linguists, phoneticians, and physicians. This was bolstered by Frank's very selective drawing upon university faculties in the northeast for part-time associates who could meet under the same roof but still meld their work at Haskins with their academic programs on campus. There was, in addition, a small full-time professional staff together with indispensable support people.

Stimulated by earlier work at Haskins on guidance devices for the blind and by ongoing work on reading machines for the blind, Franklin Cooper, Alvin M. Liberman, and Pierre Delattre were the prime movers in groundbreaking speech research, including what I call the quest for the cues to speech intelligibility. The cues being hunted were mostly those of American English, but those of other languages soon engaged the group's attention. Others who came into the group very early were Louis J. Gerstman, Katherine S. Harris, and Leigh Lisker. Frank felt strongly that an important practical outcome of this line of research would be a contribution to the generally acknowledged need for a system of synthesis of speech by rule. One pressing need for this was in the domain of reading machines. Indeed, although the practical availability of such a machine for blind users was eventually achieved by others, the feasibility of such an idea was greatly aided by Frank's projects.

I hasten to add that Frank was equally interested in the theoretical implications of this early work on speech perception. What can we learn about the transmission of linguistic information through the articulatory modulation of sound from excitation sources? Is speech perception constrained simply by auditory, that is, psychoacoustic limitations, or is there a much more abstract level that enables us to transmit and process information from speech signals at such high rates? How might children learning their first language establish boundaries in production and perception between phonologically relevant categories? Are motor control and perception separate domains, or are they intimately and systematically linked?

Work on that last question, with much momentum from the thinking of our late colleague Alvin M. Liberman, led to the famous—or in the minds of some, infamous—motor theory of speech perception. While deeply involved in research and the resulting articles on the topic, Frank Cooper never promoted a party line at Haskins Laboratories, even when one seemed to exist in the eyes of some beholders. At our frequent afternoon tea gatherings, there was much lively discussion of these theoretical issues. From those early forms of the motor theory has emerged a revision commonly known as the gestural theory of speech perception. Even so, at Haskins this appears in two versions, the one espoused by Alvin Liberman and Ignatius G. Mattingly, and the other, differing in significant ways, the one espoused by Carol Fowler. Some of the rest of us, without necessarily being advocates of one view or the other, have tried to shed light on the matter by running experiments on possible links between the production and perception of speech. Frank did not have it in his temperament to demand a party line, nor do we have one to this day.

In any event, this theoretical direction led Frank to advocate a program in electromyography to look ''upstream'' in the musculature, as he said, at the control of motor activity in speech. Perhaps there we would find the unity of the phoneme and the links with perception. Although his hopes in this regard were not realized, a great deal of important work on physiological mechanisms underlying the production of speech was carried out under the leadership of Katherine S. Harris. To this one should add the several fiberoptic studies on laryngeal and velopharyngeal behavior and x-ray motion pictures with sound stretched threefold to match the slow motion of the film.

On a number of occasions, for example, at IEEE meetings, Frank sought to convince communications engineers that these kinds of research provided an indirect but vital approach to problems of speech compression and processing. It should also be mentioned that Frank, together with other distinguished people in our field, played an important role in reporting on two trying public issues, the controversy over voiceprints and the mysterious 18-minute gap in the Nixon audio tapes.

At the instigation of the late Professor John Lotz, Frank Cooper ac-

cepted an appointment as Adjunct Professor of Linguistics at Columbia University in the fall of 1955, an unusual step for a man whose degrees were in engineering and physics. Thus he became a pioneer in introducing graduate students of linguistics, most of them with not much background in physics and advanced mathematics, to an acoustic approach to study and research in phonetics. As a member of that first class, I was indelibly marked by Frank's teaching and guidance. Each week, about seven of us took the subway train down to Grand Central Station to walk over to Haskins Laboratories in its old factory building on the three floors above a necktie factory, where we spent three or four hours listening to Frank's lectures, which were typically accompanied by revealing demonstrations. Anyone who got deeply involved in a research project was encouraged to spend extra time there.

It must be said that even in those days and before, there existed courses in speech science in departments of speech and hearing, but I believe that Frank's course in Acoustic Phonetics, named after the landmark book by Martin Joos, was one of the first, if not *the* first, to lay before students of language a new philosophy, a new outlook on how to design experiments that might answer linguistically and psychologically relevant questions about the production and perception of speech. With nothing much in print that could serve as a textbook for such a course, he led us by the hand through the reading of primary sources, the work of outstanding scientists of the day, people we have been accustomed to seeing and hearing at meetings of the Society.

In those years, a few students of psychology brought down one by one from Connecticut by Al Liberman, one from linguistics at Columbia, and one from the University of Pennsylvania did much or all of their doctoral dissertation work at Haskins with the way cleared by Frank. Indeed, for one of them, me, Frank was the principal adviser. With the move of the laboratories to New Haven, Frank, who then became Adjunct Professor of Linguistics at the University of Connecticut and Senior Research Associate in Linguistics at Yale University, put great emphasis, as soon as he could, on getting research assistantships within our grants for graduate students of Haskins staff members from those two institutions, as well as the City University of New York and the University of Pennsylvania, where we also had research associates.

Clearly, Frank Cooper was one of the luminaries in the postwar period whose impetus caused the great upsurge in research into the production and perception of speech that, happily, is still with us today. Much of the increase in this momentum should be attributed to his being a significant force in drawing into the stimulating arena of interdisciplinary research, graduate students in a variety of programs concerned with an understanding of language and speech.

I hope that I may be forgiven for ending on a personal note. Both as a scientist and an admirable human being, Frank Cooper strongly influenced my career and my life. He was my teacher, my dissertation adviser, my boss, my colleague, and my friend. I miss him very much. He is survived by his sons Alan Kent Cooper of Palo Alto, CA, and Robert Craig Cooper of Amherst, VA, four grandchildren, and two great-grandchildren.

ARTHUR S. ABRAMSON

BOOK REVIEWS

P. L. Marston
Physics Department, Washington State University, Pullman, Washington 99164

These reviews of books and other forms of information express the opinions of the individual reviewers and are not necessarily endorsed by the Editorial Board of this Journal.

Editorial Policy: *If there is a negative review, the author of the book will be given a chance to respond to the review in this section of the Journal and the reviewer will be allowed to respond to the author's comments. [See ''Book Reviews Editor's Note,'' J. Acoust. Soc. Am.* ***81****, 1651 (May 1987).]*

Fundamentals of Noise and Vibration

F. Fahy and J. Walker

E & FN Spon, London and New York, 1998.
xvii+ 518 pp. Price: $60.00.

According to the preface for this text, it ''...is based on material presented during the first semester of the postgraduate Masters Course in Sound and Vibration Studies at the Institute of Sound and Vibration Research in the University of Southampton, UK.'' At 518 pages and with a significant breadth of coverage, that first semester must be rather challenging. Nonetheless, having read this text at length, I intend to hand it to my own students with the simple directive, ''Read this.'' Paraphrasing the preface, the first semester of the ISVR Masters Course targets students with little or no prior knowledge of acoustics. Many of us find ourselves in the same situation when introducing new students to acoustics here in the U.S., and it is that application that this text serves most admirably.

The text is organized into eight chapters, with relevant references and questions at the end of most chapters. Each chapter has a different author, whose expertise and experience is appropriate to the subject matter of the chapter. Chapter 1, authored by Phil Nelson, is entitled ''An introduction to acoustics.'' The chapter introduces, in brief, many of the basic concepts in acoustics, including the wave equation, impedance, standing waves, point sources, etc. The material is presented clearly and logically, pitched at a level appropriate to students new to the field. This single chapter has a breadth comparable to some two-semester graduate courses, but without the depth. The depth is intentionally omitted here (and in subsequent chapters), yet its absence is not a flaw in the text's purpose: to support introductory studies in acoustics. Nelson includes references to a number of excellent texts where the full measure of material may be found, including Kinsler and Frey, Skudrzyk, and Pierce.

Chapter 2, authored by N. Lalor, is entitled ''Fundamentals of vibration.'' The chapter presents a straightforward development of vibration-related concepts up through two-degree-of-freedom systems. It addresses free and forced response, damped systems, transmissibility, forced excitation, isolation, etc. In addition, there is a brief development of Lagrange's equations at the end of the chapter. A weakness of this chapter is the complete lack of references to other texts with more extensive development of the theory of vibrations.

Chapter 3, authored by I. H. Flindell, is entitled ''Fundamentals of human response to sound.'' The chapter provides an excellent review of the impact of sound on humans, including such issues as annoyance, sleep disturbance, and activity interference. The chapter also addresses the physiology of the hearing mechanism, and human hearing characteristics such as masking, level dependency, etc. The chapter also addresses various measurement schemes for acoustic signals, including some limited discussion of FFT analyzers. As a minor quibble with the sequence of information within the text, frequency analysis is not introduced until Chap. 6, and as there is nothing in this section to direct the reader to Chap. 6 for more detail, the FFT material here seems a bit out of place. (I would have preferred the chapter on frequency analysis to have immediately followed Chap. 2.)

Chapter 4, authored by M. J. Griffin, is entitled ''Fundamentals of human responses to vibration.'' This material is frequently absent from other introductory acoustics texts, and even from most texts on vibrations, so its inclusion here represents a unique resource. The chapter covers measurement systems, health impacts, and recommended practices for human exposure to vibration. As one might imagine from a text originating in the UK, many of the impact studies and standards that are referenced in this chapter (and in the other chapters) are of British or European origin. This does make some of the material of less relevance to U.S. readers and practitioners, but not unduly so.

Chapter 5, authored by F. J. Fahy, is entitled ''Fundamentals of noise and vibration control.'' The chapter begins with a brief review of some basic acoustics terminology (duplicating material found elsewhere), and then moves into a quite thorough and valuable exposition on the nature and characterization of noise sources. The preliminaries out of the way, the chapter then proceeds to review general concepts in passive noise control, including source–path–receiver considerations. The chapter then presents specific theory on sound absorption, transmission through partitions, and barriers. The chapter concludes with a brief review of vibration control concepts. On the whole, the chapter is a good introduction to the fundamental concepts and a few of the more commonly encountered theoretical matters. The references direct the readers to more comprehensive treatises on noise control.

Chapter 6, authored by J. K. Hammond, is entitled ''Fundamentals of signal processing.'' The chapter presents a broad overview of the theoretical basis for digital signal processing. It concisely covers many of the topics that first-time students find so confounding (or of which they are blissfully ignorant) when using the now ubiquitous FFT signal analyzers, whether dedicated or PC based. Such topics include sampling effects and windowing, and application of various statistical measures. The chapter also presents the topic of system identification. My sole issue with this chapter is one of sequence: given that other preceding chapters had reason to resort to discussions of frequency analysis and analyzers, this chapter more logically should have followed the second chapter.

Chapter 7, authored by T. G. Leighton, is entitled ''Fundamentals of underwater acoustics.'' The chapter opens with a broad discussion of matters related to sound propagation in the oceans, including sound speed profiles, influence of bubbles, and ray acoustics. The chapter then discusses a variety of nonlinear effects in underwater acoustics, including finite-amplitude propagation, self-effects, and parametric arrays. The chapter concludes with a brief discussion of the medical effects of ultrasound. While oddly juxtaposed with a chapter that otherwise deals with sound propagation in the oceans, the inclusion of the ultrasound material here is probably due to ISVR's program in medical acoustics.

Chapter 8, authored by R. J. Pinnington, is entitled ''Fundamental principles of measurements and analysis techniques.'' This chapter is a fitting conclusion to the text, as it presents the issues and theory related to the measurement of acoustic and vibration signals, the topics of the previous chapters. For novice students in experimental programs, this chapter will be quite an eye-opener as it describes the interactions between exciter and test articles and frequency limitations of devices. The chapter concludes with a very brief discussion of practical issues in measurement, such as noise, cross-talk, and digitization effects. However, the utility of this concluding section is limited by its brevity and its lack of references to more detailed expositions on these very important issues.

Overall, the text is well organized and presented. Most chapters have good references and some sample problems. A solutions manual is available. The depth of the material in each chapter is admittedly limited, yet its

 0001-4966/2000/108(5)/1972/4/$17.00

absence is not a flaw in the text's purpose: to support introductory studies in acoustics. The preface states that a companion volume based on the second semester course at ISVR, as well as their Advanced Course in Noise and Vibration will follow: I look forward to reading that text, as well.

KENNETH A. CUNEFARE
Georgia Institute of Technology
G. W. Woodruff School of Mechanical Engineering
Atlanta, Georgia 30332-0405

The Boundary Element Method in Acoustics

Stephen Kirkup

Integrated Sound Software, West Yorkshire, UK, 1998.
xii+135 pp. Price: $63.00 (Core Package: Book and ABEM2D disk) and $499 (Development Package: Book and ABEMFULL CD).

The subject of the boundary element method (BEM) in acoustics has been discussed by so many people over the past decade that it no longer draws any excitement or attention to me. I have collected at least one book (Ciskowski and Brebbia, 1991) that bears the same title as the current one. So my first reaction to the request of review was why another book on BEM? It turned out that this book is concerned with the implementation of BEM using FORTRAN 77 for acoustic radiation and scattering in both interior and exterior regions. The computer program is known as Integrated Sound Software (ISS), and the BEM codes are written in the simplest forms to allow for flexibility in engineering applications. Specifically, the boundaries are approximated by straight lines for two-dimensional problems, planar triangles for three-dimensional problems, and truncated cones for axisymmetric three-dimensional problems. The functions are approximated by constants or C^0 interpolation on the boundary elements. The non-uniqueness difficulties inherent in the Helmholtz integral theory for the exterior region are overcome by Burton and Miller's method (1971). Tests on various subroutines and the corresponding results are demonstrated in Chaps. 4, 5, and 6 for interior and exterior acoustic problems, and interior modal analysis, respectively.

Overall this book serves as an extended user manual for the BEM codes to solve acoustic radiation and scattering problems in interior and exterior regions, with a brief discussion of the Helmholtz integral theory and listing of formulations for different situations. As such it seems balanced. However, I have the following comments on the subject matter.

(1) While the integral on the right side of (3.37) is finite, the first two terms are not in the limit as $q \to p$. Detailed justifications must be given in order for this general relationship to hold. My guess is that (3.37) holds simply because a planar boundary element is used.
(2) It is not clear how the singularities are handled as the field point approaches the surface in this book. Figure 3.1 shows a diagram in which the planar triangle element is divided into three parts when $p \in \Delta\tilde{\Gamma}$. However, exactly how this division will circumvent the singularity is not discussed.
(3) Discussions on scattering in an interior region on p. 60 are meaningless. Unlike scattering in an exterior region, the incident wave in an interior region is not well defined because of the presence of reflected waves. Hence it is inappropriate to copy the treatments of scattering for an exterior region to an interior region.
(4) The field solution inside a sphere of unit radius should not be $\varphi(p) = e^{ikr}/r$ (see p. 75), because it is unbounded when $r=0$. Justifications are needed here.

The computer codes discussed in this book were also examined, which was done as a part of class projects for ME7460: Advanced Acoustic Radiation that I taught at Mechanical Engineering Department of Wayne State University during the Winter semester, 2000. The main topics covered in this course included the Helmholtz integral theory and its implementation using BEM, and near-field acoustic holography. The enrollment of this class was comprised of three Ph.D. and eight MS students. The students taking this course were all very highly motivated and hard working. One of the projects was to use ISS and SYSNOISE® (another commercial software developed by Leuven Measurement Systems) to simulate acoustic radiation from a dilating sphere, an oscillating sphere, and an arbitrarily shaped object. The integral theory and its numerical implementation were discussed at length in class. The students were required to learn ISS and SYSNOISE to solve problems on their own. At the end of the semester, students were asked to fill out surveys of the book and software ISS, which consisted of ten questions. The first five questions were about the clarity of the book. Question 6 was a self-assessment of learning ISS. Question 7 was concerned with the interface of ISS with other software in generating boundary elements and meshes. The last three questions were concerned with the students' recommendation and

TABLE I. Responses to questions 1–8.

No.	Questions	Absolutely	Moderately	Very little	Not at all
1	Have you read the book entitled *The Boundary Element Method in Acoustics*?	36%	64%	0%	0%
2	Do you feel that this book has adequately described the Helmholtz integral theory?	18%	64%	9%	9%
3	Do you feel that this book has adequately described the boundary element method?	36%	46%	18%	0%
4	Do you feel that this book has adequately explained how to run the Integrated Sound Software?	9%	37%	27%	27%
5	Do you feel this book has provided enough examples for you to understand the software for different cases?	18%	18%	37%	27%
6	Do you feel that you can learn how to run Integrated Sound Software on your own?	27%	18%	45%	10%
7	Do you feel that the interface of Integrated Sound Software is straightforward and user-friendly?	0%	0%	18%	82%
8	Would you recommend Integrated Sound Software to your friends?	27%	18%	37%	18%

TABLE II. Responses to questions 9 and 10.

No.	Questions	Excellent	Very good	Good	Fair	Poor
9	How would you rate the book?	9%	9%	36%	18%	28%
10	How would you rate the Integrated Sound Software?	0%	36%	10%	27%	27%

their ratings of the book and software. The results of this survey are shown in Tables I and II.

Most of the responses to the questions were mixed. The overall feeling for the software was not very encouraging, however. The biggest problem with ISS was its interface with other software in generating the boundary elements or meshes. Every student complained bitterly about the interface of this software. On the other hand, the core program can produce relatively good results, even though constant values or a C^0 interpolation scheme is used to approximate the functions on the boundary surface.

Based on the students' feedback and my personal review, I must conclude that this book falls short of providing general BEM codes that can be handy for people working in the field of acoustic radiation and scattering. I strongly recommend that the author consider adding a component that may make the interface of ISS with other software easier in generating boundary element meshes, thus making ISS more accessible to users.

SEAN F. WU
Acoustics, Vibration, and Noise Control Laboratory
Department of Mechanical Engineering
Wayne State University
Detroit, Michigan 48202

Comparative Hearing: Insects

Ronald R. Hoy, Arthur N. Popper, and Richard R. Fay, eds.

Springer-Verlag, New York, 1998.
341 pp. Price: $89.95 (hardcover) ISBN: 0387946829.

This is the tenth in the series of the Springer Handbook of Auditory Research edited by Fay and Popper. The lead editor for this volume is Dr. Hoy, a well-known neurobiologist in insect acoustics. In addition to Dr. Hoy, seven other acknowledged experts have contributed to this interesting book.

In the introductory chapter by Hoy, ''Acute as a Bug's Ear: An Informal Discussion of Hearing in Insects,'' it is indicated that the emphasis throughout the book is on tympanal hearing of airborne sound pressure. This limits the scope of the book. Aquatic insects are of course excluded. A more important omission (except for a few brief comments in Chapter 2 and elsewhere) is the group of insects that sense the particle velocity of sound with hair sensilla or antennae in the near field. The emphasis in the book is thus on sensing in the far field. Dr. Hoy provides a readily understood introduction to the tympanal hearing systems of insects. It is indicated that all insects, including the praying mantis, have two ears. In contrast to the ears of vertebrates that are located in the cranium, ears can be found nearly anywhere on an insect's body. How and why this happens is clearly explained. Some details of insect hearing are mentioned briefly, including the fact that insects can be quite sensitive to the direction of a sound source. Directional hearing is a recurring theme in the book. Also, the relatively short life span of insects makes it easier to investigate the evolution of insect hearing.

The second chapter, ''Biophysics of Sound Localization in Insects'' by Michelsen, provides a summary of basic principles, including a listing of three mechanisms of acoustic reception: pressure receivers, pressure-difference receivers, and particle-motion detectors. Pressure-difference receivers and particle-motion detectors are inherently directional. However, the distinction between the two is not clear-cut because pressure-difference receivers can be used to measure particle motion. Directional hearing is required by insects, principally to avoid predators and for mating. The directional hearing of humans and other vertebrates can be explained on the basis of diffraction around the cranial structure and distance between the ears. In a similar way, diffraction and distance between the ears on the thorax of night-flying moths can be used to determine the direction of the ultrasonic signals of bats. In general, however, insects are too small relative to wavelengths of interest to be able to use diffraction and distance between the ears. Thus insect ears generally cannot be considered to function purely as pressure receivers. To obtain directionality, the tympanal structure of the ear has to act as a kind of pressure-difference receiver with pressure exerted by the external sound field on the outside and, on the inside, pressure conveyed through tracheal passages or other air-filled tubes, responding to input from the other ear. This explanation has been worked out in detail for a few crickets and certain grasshoppers. Unlike Hoy, Michelsen states that the praying mantis has only one ear.

In Chapter 3, ''The Sensory Ecology of Acoustic Communication in Insects,'' Roemer is concerned with natural effects that attenuate the transmission of insect signals in the far field, such as absorption in the atmosphere, scattering by vegetation and atmospheric turbulence, the effect of temperature and humidity gradients, and transmission over the ground. These can cause significant frequency-dependent losses in signal strength. To counter such effects, insects have developed ingenious methods of improving acoustic communication, such as the mole cricket amplifying its signals by matching the dimensions of its burrow to the frequency of its call and the South African tree cricket using a leaf as an acoustic baffle. Another obvious method is the redundancy of repetitious, broadband, stereotyped signals. In ways such as these, insects can extend the range and improve the detection of signals to a remarkable degree.

Boyan in Chapter 4 discusses the development of tympanic auditory systems in insects. This is illustrated for three insect groups: grasshoppers, bushcrickets (katydids), and crickets. Section 4 of Chapter 1 identifies these insects and their hearing systems. Development of the auditory system is traced through the embryonic and post-embryonic stages. Only at the adult stage is the insect ear fully functional. Information about the development of the auditory system is also provided by regeneration experiments in which certain parts of the ear can grow again after being removed. The studies described in this chapter provide important clues about how insect ears evolved.

In Chapter 5 Pollack discusses the neural processing of acoustic signals by insects. The first sections summarize the characteristics of sound signals most apparent to insects and subsequent sections describe how these features are detected and analyzed by the nervous system. The material relates principally to the three insect groups used as examples in Chapter 4.

In Chapter 6, ''The Evolutionary Innovation of Tympanal Hearing in Diptera,'' Robert and Hoy consider the directional hearing of a parasitoid fly that is much smaller than its cricket host. A pregnant ormiine fly deposits a few first-instar larvae on or around a field cricket which then enter the cricket by themselves. After about a week, third-instar larvae emerge from the cricket and pupate outside. It is known that the fly detects the cricket host by homing in on the cricket's song and for this purpose it is endowed with keen directional hearing. The song, of course, is not meant for the parasitoid fly but for other field crickets. Because of its much smaller size, the fly's directional hearing has to be different from a cricket. It is this novel hearing capability that is the subject of the chapter. Essentially the very small differences in the time of arrival of the acoustic signals at the two auditory tympana of the fly are mechanically amplified by viscoelastic coupling. Neural processing then provides additional amplification needed for good directional hearing. This work is important because it provides fresh insights into the evolution of directional hearing in insects.

In Chapter 7, ''The Vibrational Sense of Spiders,'' Barth discusses substrate transmission of acoustic signals. It might be asked why the book has a chapter on spiders when there are known examples of insects using substrate transmission. The biomechanics of the transmission process is the same for both, however. As well as discussing transmission paths that can apply to insects, for example in the leaves and stems of plants, Barth provides an interesting discussion of transmission across webs and water surfaces. Spiders are important because they combine, to a high degree, detection of substrate vibration with detection of air movement using hair sensilla. Barth reviews the mechanics of vibrations in different types of substrates and describes vibration receptors in spiders, the most prominent and sensitive being the metatarsal lyriform organ. How spiders generate vibrations is discussed, specifically for *Cupiennius getazi*, and lycosids (wolf spiders) that combine drumming on the substrate with stridulation. This is followed by a review of vibration signals used by spiders and the kinds of sound attenuation and interference encountered in their environment.

In the final chapter, ''The Sensory Coevolution of Moths and Bats,'' Fullard discusses the classic bat–moth story. Man has been intrigued by the ability of bats to sense in the dark and by the ability of moths in flight to avoid bat predators. However, full understanding of the bat–moth interaction was achieved only in the middle of the last century, as a result of the

work of Griffin on the ultrasonic signals of bats, followed by the moth studies of Roeder. Fullard recounts the fascinating story of the evolution of the interaction and how it is continuing to evolve.

The book covers topics that are fairly well known to the average person: the sounds of crickets, the behavior of spiders in their web, and the interaction between bats and moths. Less well known, but nevertheless important topics, such as the use of hair sensilla to sense the particle velocity of sound in the new field and the sounds of aquatic insects, are not covered. Perhaps there is not enough definitive information, but these topics are a significant part of insect hearing and should have been included. The sounds of aquatic insects could be an indicator of the ecological health of fresh water systems, such as rivers, lakes, and wetlands.

I found the book informative and stimulating. Scientists involved with hearing should find the book rewarding. Hopefully it will also interest other kinds of acousticians, perhaps encouraging them to contribute to the field of animal acoustics. If we fully understood the capabilities of bats, the whole science of acoustics would be greatly advanced from where it is now.

ROBERT HICKLING
Sonometrics, Inc.
8306 Huntington Road
Huntington Woods, Michigan 48070

Speech Processing and Synthesis Toolboxes

Donald G. Childers

John Wiley & Sons, New York, 2000.
483 pp. Price: $90.95 (hardcover) ISBN: 0471349593.

The book *Speech Processing and Synthesis Toolboxes* by Donald Childers is a valuable addition to the scarce literature on speech analysis and synthesis. It provides several software toolboxes that are not available elsewhere. An extensive database of speech files also accompanies the book, and the theory behind the software is explained. It is the result of many years of experience and work in this area by the author.

This book begins with a chapter titled ''Software: Installation and Introduction,'' which contains a clear description on the differences between the two software versions provided (the runtime server version and the regular MATLAB version). I believe it was a great idea to provide the runtime server version, since many of the potential users may not have MATLAB, which is required to run the regular version. I had some difficulty installing the software, in particular, I could not find the formtk_4 files nor the Chap_10 folder.

Chapter 1 is an introduction to the book. It begins by telling the background assumed for the material covered in the text, and also presents some of the applications in speech analysis and synthesis, which I found very motivating. An introduction and explanation of some of the terminology commonly used in speech, such as fricatives, pitch, formants, etc., is given. The first software toolbox (speech_1_display) is used and demonstrated. Chapter 2 is called ''Speech Analysis Toolbox.'' It mainly shows how to use the Speech Analysis Toolbox by guiding the user through examples. It does not explain the theory behind the software since it is explained in later chapters and Appendix 6. I was very impressed by all the different types of analysis that can be done with this toolbox (time, frequency, glottal inverse filtering, pitch/jitter/formant contours, cepstral, and WRLS-VFF). This toolbox provides windows for the different types of analysis so that the user can vary the parameters at will and observe the effects on the results.

''Speech Production, Labeling and Characteristics'' is a theoretical chapter devoted to the description of general and specific aspects of speech production and its classification. Many figures and tables are provided to help understand the theory. It becomes evident in this chapter why the author stated in Sec. 1.3 that the text assumes that the speech production and characterization is for American English, which is important for potential foreign buyers. The ''Data and Measurements'' chapter describes a series of methods used in speech to obtain information about the laryngeal function. It is obvious from this chapter that there is a wide range of measurement procedures. The difficulty in directly measuring the voice source can also be seen. A toolbox is used to display and analyze speech and EGG (electroglottographic) signals.

Chapter 5 is called ''Linear Prediction.'' In this chapter the author presents the theory about the most commonly used spectral estimators, which are classified as classical and parametric estimators. This is not an easy subject, since the theory on parametric estimators is complex, and there is disagreement among authors about some concepts like zero padding and windowing; however, the author presents the concepts in a clear and precise way. ''Speech Synthesis and a Formant Speech Synthesis Toolbox'' describes the most commonly used methods for speech synthesis (Articulatory Synthesis, Formant Synthesis, and LP Synthesis), and a toolbox for speech synthesis. The use of the toolbox helps the user understand the effect of the different parameters involved in creating synthesized speech, and the steps needed to generate it.

The ''VOCOS-A Voice Conversion Toolbox'' chapter mainly describes how to use this toolbox, which provides a relatively large amount of options that can be summarized in three groups: analysis of a speech file, modification of the file parameters, and speech synthesis. The reader may not be familiar with some of the terminology used in this chapter and may need to read Appendices 6, 7, and 8 to obtain additional information. Also, the user will realize the large amount of parameters associated with a speech file while using VOCOS. Chapter 8 is called ''Time Modification of Speech Toolbox.'' This chapter shows the options of this toolbox, which consists of three main functions: (1) The analysis option, which segments the speech file into phonemes and labels them; (2) the modify option, which allows the user to change the segment boundaries and their labels; (3) the user can specify the time modification parameters for each phoneme and create the synthesized data using the time modification and synthesis option. The quality of the synthesized sample I created with this toolbox was very good.

Chapter 9 describes the ''Animated Vocal Fold Model Toolbox.'' This toolbox contains two vocal models: the Two Mass Model, and the Ribbon Model. Both of them simulate the vocal fold motion. The toolbox is very simple to use and the animated model is created in three steps: generation of the glottal area waveform, drawing of the prephonatory shape, and creation of the vocal folds animation. The main functions of the ''Articulatory Speech Synthesis Toolbox'' are described in Chap. 10, which can be summarized as follows: (1) The first four formant tracks are extracted from a target speech file; (2) the initial shape of the vocal tract is specified through a set of parameters; (3) an articulatory parameter vector is determined by applying inverse filtering; (4) an excitation waveform is designed; (5) the synthesized speech is created. The five steps previously described require that the user set a relatively high number of parameters, and the synthesis process is computationally intense (it takes 1.5 h approximately to synthesize the word ''be'' using 13 frames), which shows the user the difficulty in modeling the vocal tract. At the end of the chapter there is a list of synthesized files provided with the software. The book has 13 Appendices, which contain a glossary of terms, references, standards, the theory behind some of the toolboxes, and a list of the toolboxes.

During the entire software testing phase I was using MATLAB version 5.3. Although the software was tested by the author with version 5.2, I only had minor problems using version 5.3. Whenever I encountered problems, I used the runtime server version, which ran properly at all times. I believe the author's claim ''to strike a balance between theory and practice'' was fulfilled. The software provided contains many functions that would be too long for the reader to develop, but are important for the understanding of the theory on speech analysis and synthesis. In summary, this is a great book that will be an invaluable tool for anyone working in the speech analysis and synthesis field. The author has made an enormous effort in gathering a big amount of information in a single book. Foreign buyers must be aware that the speech production and characterization is for American English.

JOSE A. DIAZ
Universidad de Carabobo
Valencia, Venezuela

REVIEWS OF ACOUSTICAL PATENTS

Lloyd Rice
11222 Flatiron Drive, Lafayette, Colorado 80026

The purpose of these acoustical patent reviews is to provide enough information for a Journal reader to decide whether to seek more information from the patent itself. Any opinions expressed here are those of reviewers as individuals and are not legal opinions. Printed copies of United States Patents may be ordered at $3.00 each from the Commissioner of Patents and Trademarks, Washington, DC 20231.
Reviews of Acoustical Patents

Reviewers for this issue:

GEORGE L. AUGSPURGER, *Perception, Incorporated, Box 39536, Los Angeles, California 90039*
JOHN ERDREICH, *Ostergaard Acoustical Assoc., 200 Executive Drive, West Orange, New Jersey 07052*
DAVID PREVES, *Songbird Medical, Inc., 5 Cedar Brook Drive, Cranbury, New Jersey 08512*
CARL J. ROSENBERG, *Acentech, Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*
ERIC E. UNGAR, *Acentech, Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*

6,043,924

43.35.Sx CONSTANT POWER DENSITY ACOUSTIC TRAVELING WAVE LENS

Robert M. Montgomery *et al.*, assignors to Harris Corporation
28 March 2000 (Class 359/285); filed 22 January 1999

This patent relates to laser scanning systems, in which scanning is accomplished by passing a laser beam through an acoustic traveling wave lens. This lens consists of a fluid-filled channel, which serves as an acoustic waveguide with an exciting transducer at one end and a termination on the other. A waveguide according to this patent is tapered, so that it becomes narrower with increasing distance from the exciter, thereby maintaining constant acoustic power density in the fluid medium by compensating for acoustic attenuation. A system according to this patent also includes a heater to maintain the medium at a temperature at which the acoustic velocity is relatively independent of changes in temperature.—EEU

6,056,696

43.35.Sx FRUSTRATED TOTAL INTERNAL REFLECTION ACOUSTIC FIELD SENSOR

Jeffrey S. Kallman, assignor to The Regents of the University of California
2 May 2000 (Class 600/459); filed 4 September 1998

This patent is intended to provide means for observing the acoustic field over an entire plane at once, thus facilitating use of this information for acoustic holography and acoustic diffraction tomography. The basic principle employed here is that of frustrated total internal reflection of light, which is explained as follows. As light encounters the surface of a medium in which the speed of light is greater, total reflection occurs if the angle of incidence is greater than a critical value. However, in this situation there also occurs an evanescent wave that extends into the ''fast'' medium. If another ''slow'' medium is placed beyond the fast medium, this second slow medium intercepts the evanescent wave and some light tunnels through the gap formed by the fast medium between the two slow media. The amount of light that tunnels is directly related to the gap width.

The sensor described in this patent in essence consists of a transparent plate, near which a flexible membrane is placed at a distance suitable for the occurrence of frustrated total internal reflection. The membrane is in contact with a liquid ultrasound medium on the side away from the transparent plate, so that an acoustic wave in the ultrasound medium causes the membrane to deflect, thus changing the gap between it and the transparent plate. When a uniform light beam is directed through the transparent plate at a suitable angle, some of the light tunnels through gap. The rest is reflected and sensed to provide a measure of the membrane's deflection and thus of the field in the ultrasonic medium.—EEU

6,058,076

43.35.Yb SIGNAL DETECTOR AND METHOD FOR DETECTING SIGNALS HAVING SELECTED FREQUENCY CHARACTERISTICS

Nikolaos I. Komninos, Littleton, Colorado
2 May 2000 (Class 367/135); filed 8 December 1997

The subject of this patent is related to the detection of leaks in pipelines, of small openings in pressurized buildings or other volumes, and of wear in bearings and the like. An embodiment employs an ultrasonic microphone, associated filters to remove frequency components that are not of concern, and circuitry to mix the resulting signal with a reference signal. In addition, it uses a contact probe to sense acoustic emissions, also with associated filters and mixing circuitry. Either of the resulting outputs may be displayed visually.—EEU

6,014,609

43.38.Gn ACOUSTIC REFLECTION CHART RECORDER

James N. McCoy, Wichita Falls, Texas
11 January 2000 (Class 702/6); filed 1 April 1998

Echo ranging is a principal technique for measuring the location of the liquid surface of fluid in an oil well. Sound reflections generated from impulse excitation are picked up by a microphone and analyzed in relation to known locations of tubing collars. The patent describes a more sophisticated, automated variation of this technique which includes first monitoring background noise in the well bore to establish maximum pickup gain, and then automatically varying gain in response to signal level.—GLA

6,015,025

43.38.Ja DIFFUSER PANEL WITH BUILT-IN SPEAKER ARRANGEMENT AND METHODS OF INSTALLATION

Ralph D. McGrath and Harry A. Alter, assignors to Owens Corning Fiberglass Technology, Incorporated
18 January 2000 (Class 181/199); filed 6 June 1997

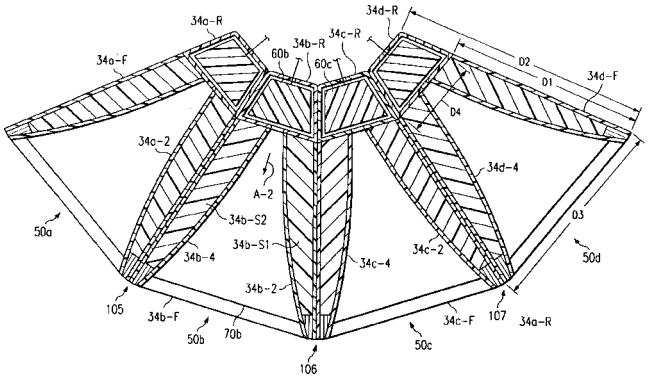
This long-winded patent can be summarized in three sentences: A room used for home theatre requires loudspeakers. Room acoustics can be improved by adding absorption and diffusion. Why not combine all three?—GLA

6,016,353

43.38.Ja LARGE SCALE SOUND REPRODUCTION SYSTEM HAVING CROSS-CABINET HORIZONTAL ARRAY OF HORN ELEMENTS

David W. Gunness, assignor to Eastern Acoustic Works, Incorporated
18 January 2000 (Class 381/342); filed 29 August 1997

The illustration is a horizontal section through a compact, practical four-horn array. However, trapezoidal cabinets that can be bolted together to form arrays have been available from a number of manufacturers (including



the assignee) for many years. Apparently, no one thought to patent an array of trapezoidal cabinets bolted together. To a patent attorney, the difference is no doubt profound.—GLA

6,031,926

43.38.Ja PANEL-FORM LOUDSPEAKERS

Henry Azima *et al.*, assignors to New Transducers Limited
29 February 2000 (Class 381/423); filed 2 September 1996

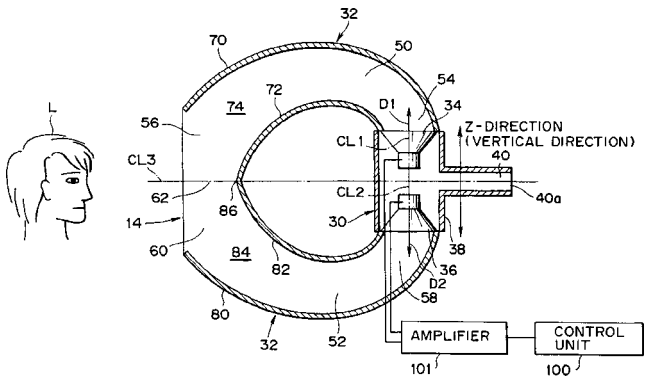
The patent abstract, in its entirety, reads, ''A panel-form loudspeaker has two distributed mode acoustic radiators coupled by a resilient suspension. Each radiator has its own transducer.''—GLA

6,035,051

43.38.Ja SOUND APPARATUS

Masaru Sato, assignor to Sony Corporation
7 March 2000 (Class 381/340); filed in Japan 12 May 1997

To make use of available space in stereo TV receivers, loudspeakers can be mounted in the rear of the cabinet and sound conducted to the front through waveguides. In this variant, two vertically symmetrical waveguides **74** and **84** cross-fire through a common aperture **14**. Accordingly, ''...sounds



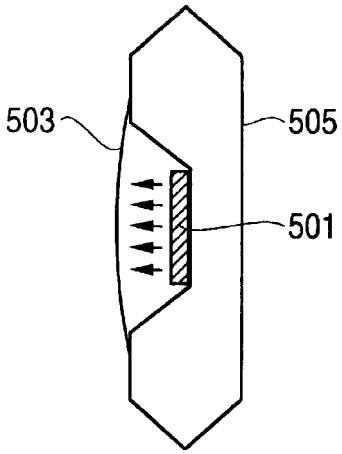
having the same phase are synthesized by collision... .'' It is rare to find a Sony patent that seems to make no sense at all, even allowing for translation from the original Japanese.—GLA

5,949,892

43.38.Kb METHOD AND APPARATUS FOR DYNAMICALLY CONTROLLING OPERATING CHARACTERISTICS OF A MICROPHONE

Brett B. Stewart, assignor to Advanced Micro Devices, Incorporated
7 September 1999 (Class 381/113); filed 7 December 1995

This novel micromachined microphone includes a charge-carrying circuit **501** between substrate **505** and diaphragm **503**. An applied charge con-



trols the compliance of the diaphragm. Operating characteristics such as dynamic range and frequency response can thus be adjusted dynamically.—GLA

6,038,325

43.38.Lc SPEAKER SYSTEM FOR USE IN AN AUTOMOBILE VEHICLE

Minoru Yoshida *et al.*, assignors to Poineer Electronic Corporation
14 March 2000 (Class 381/86); filed in Japan 6 February 1997

Auto passenger compartment geometry is such that the lowest resonance usually produces a response peak in the rear seat, but a semi-null in the front seat. The patent teaches that response can be spatially equalized by introducing a rear speaker as an active acoustic element. The speaker is driven by an L+R signal processed by a bandpass filter and phase-shifting network. This is a clever idea.—GLA

5,954,630

43.38.Lc FM THETA-INDUCING AUDIBLE SOUND, AND METHOD, DEVICE AND RECORDED MEDIUM TO GENERATE THE SAME

Kazumi Masaki and Osamu Matsuda, assignors to Ken Hayashibara
21 September 1999 (Class 600/28); filed in Japan 16 September 1993

This must be the first acoustical patent to include photos of brain scans. It can be argued that, since theta-type brain waves are associated with attention and concentration, then artificially induced theta waves should improve these mental qualities. Experiments by the inventors suggest that an

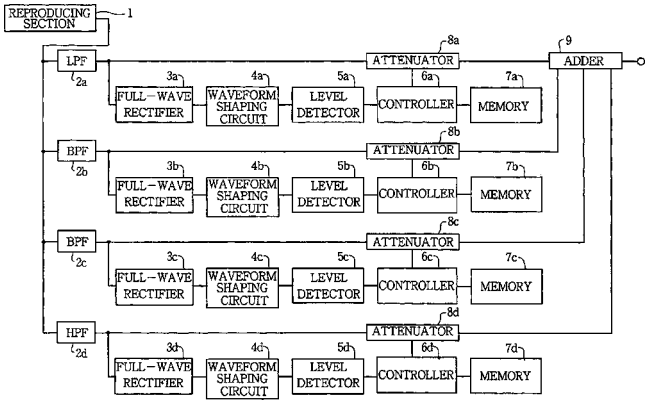
effective stimulus is a low-frequency audible sound (100–500 Hz) amplitude modulated at 20 Hz or lower. Practical embodiments are described in the patent.—GLA

5,926,334

43.38.Md NOISE REDUCTION SYSTEM FOR AN AUDIO SYSTEM

Shirou Suzuki, assignor to Pioneer Electronic Corporation
20 July 1999 (Class 360/55); filed in Japan 9 May 1996

This single-ended noise reduction system dynamically controls relative levels of individual frequency bands. It is intended to suppress tape hiss. Only one reference is cited, a 1986 patent. One would expect at least some



mention of the wealth of noise reduction circuits (scratch supressors) developed in the 1940s and 1950s.—GLA

5,930,370

43.38.Vk IN-HOME THEATER SURROUND SOUND SPEAKER SYSTEM

Jerome E. Ruzicka, assignor to REP Investment Limited Liability
27 July 1999 (Class 381/18); filed 7 September 1995

Outputs from left and right power amplifiers are translated into four-speaker surround sound with almost no additional circuitry. Left and right speakers are high-pass filtered. A dual-coil center speaker combines left plus right signals. A rear speaker reproduces L–R (or R–L). The inclusion of high-pass filtering seems to be the one novel feature of this invention.—GLA

6,016,295

43.38.Vk AUDIO SYSTEM WHICH NOT ONLY ENABLES THE APPLICATION OF THE SURROUND SYSTEM STANDARD TO SPECIAL PLAYBACK USES BUT ALSO EASILY MAINTAINS COMPATIBILITY WITH A SURROUND SYSTEM

Kenjiro Endoh and Craig Campbell Todd, assignors to Kabushiki Kaisha Toshiba
18 January 2000 (Class 369/49); filed in Japan 2 August 1995

The patent document consists mostly of computer flow diagrams. A bit stream signal not only contains audio information but also a mode flag that determines how the audio signals are to be decoded. Thus, what would normally be surround sound audio might become something quite different when a special mode such as karaoke is called for.—GLA

6,038,324

43.38.Vk AUTOMOTIVE SURROUND SOUND CIRCUIT BACKGROUND OF THE INVENTION

Paul R. Ambourn, North Oaks, Minnesota
14 March 2000 (Class 381/18); filed 21 February 1997

The patent title clearly refers to automotive sound. All six claims are specifically limited to automotive surround sound. However, the first sentence of the abstract begins, ''A surround sound converter for home audio systems...'' and the patent text repeatedly refers to home audio systems. Such are the subtleties of the U.S. patent process.

Known prior art synthesizes a surround channel by connecting a rear speaker between positive left and right outputs. The derived signal can be either L–R or R–L depending on the polarity of the speaker terminals. But why agonize? Use two speakers instead. Be warned, however, that the connections shown in the patent document will not work if any two of the four rear output terminals share a common connection.—GLA

6,050,084

43.40.Gf SOUND ABSORBING FLUE-GAS DUCT FOR A GAS AND STEAM TURBINE PLANT

Herbert Schutz *et al.*, assignors to Siemens Aktiengesellschaft
18 April 2000 (Class 60/39.5); filed in Germany 5 September 1996

The muffling arrangement described in this patent in essence consists of a series of parallel sound-absorbing baffles that are located near the sound source, where the flow velocities are high and the flow is irregular. In order to smooth the flow over the panels and thus to reduce the mechanical loading on them, the leading edge of each panel is provided with a bulbous ''nose.''—EEU

5,744,715

43.40.Ng TACTILE-ACOUSTIC INFORMATION DETECTING AND MEASURING APPARATUS AND METHOD

Jean-Paul Gerome, assignor to Apollon's Algebra (Gibralter) Limited
28 April 1998 (Class 73/432.1); filed 5 August 1994

This device seems to use a microphone to transduce vibration produced by rubbing a membrane mounted above it. The inventor also seems to include a tuneable device to simultaneously measure vibrations produced.—JE

6,001,073

43.40.Ng DEVICE FOR INDUCING ALTERNATING TACTILE STIMULATIONS

Jurgen G. Schmidt and Shirley Jean Schmidt, San Antonio, Texas
14 December 1999 (Class 601/72); filed 22 July 1997

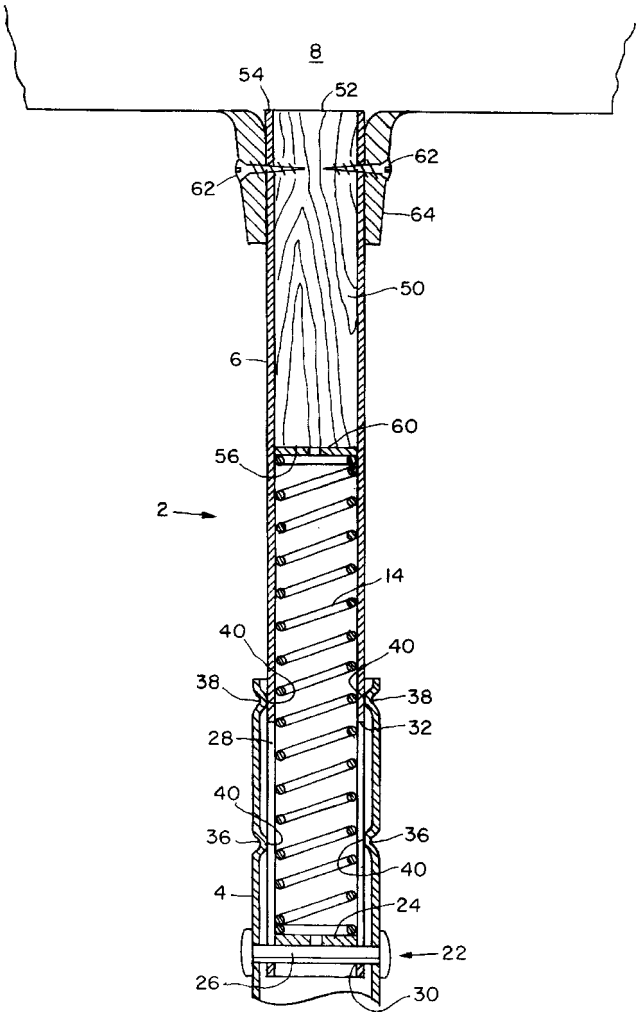
This device develops alternating bursts of vibrotactile stimulation at two hand-held transducers. It is claimed to be useful in the treatment of post-traumatic stress syndrome by a technique known as eye movement desensitization and reprocessing.—JE

5,816,634

43.40.Tm SHOCK-ABSORBING TOOL HANDLE

Paul J. Jacobs and Ronald J. Shaud, assignors to Nor-Easter Enterprises, Incorporated
6 October 1998 (Class 294/57); filed 19 April 1996

A shock-absorbing tool handle in which impacts are dissipated by a



spring mounted in a tubular arrangement mounted between the proximal and distal portions of the handle.—JE

5,929,395

43.40.Tm VIBRATIONAL ENERGY ABSORPTION PLATFORM

F. Peter Bizlewicz, Wayne, New Jersey
27 July 1999 (Class 181/207); filed 29 December 1997

This isolator sandwich is placed between an electronic component and its support structure to isolate vibration that originates in the electronic component.—CJR

5,939,179

43.40.Tm CONSTRAINT TYPE VIBRATION DAMPING MATERIAL

Kunihiko Yano *et al.*, assignors to Nichias Corporation
17 August 1999 (Class 428/212); filed in Japan 29 March 1995

Constrained-layer viscoelastic damping configurations, particularly for use as shims in brakes in order to prevent brake squeal, are made by coating a rubber layer on a metal base plate and depositing a hard thermosetting plastic atop the rubber layer. Constrained-layer damping has been well known for at least four decades, but the technology described in this patent is claimed to be inexpensive and effective.—EEU

5,971,347

43.40.Tm VIBRATION DAMPER

Chong-Shien Tsai, Williamsville, New York and Lap Lai Chung, Taipei City, Taiwan
26 October 1999 (Class 248/618); filed 24 June 1998

This isolation device appears to be conceived to protect a mounted item from shock or vibration acting in a specific horizontal direction; the relevant details are not spelled out in the patent. The device consists of two parallel horizontal plates that are interconnected by tapered flexure elements.—EEU

6,053,269

43.40.Vn VEHICLE/BRIDGE VIBRATION MITIGATION ASSEMBLY

William N. Patten, assignor to The Board of Regents of the University of Oklahoma
25 April 2000 (Class 180/167); filed 13 August 1998

Based on the observation that the natural frequencies of truck suspension systems fall into the range in which many highway bridges have natural frequencies, this patent delineates a means for avoiding resonant excitation of bridges by trucks that traverse them. An actuator installed in parallel with a truck's suspension springs in essence adjusts the suspension's stiffness and damping in response to motion data. Such data may be derived from sensors on the vehicle, with control based on the dynamics of the vehicle chassis and on estimates of the bridge motion. Alternatively, the controller also may receive data on the bridge motion via wireless communication. In other realizations, a control system may be included that acts directly on the bridge.—EEU

5,927,686

43.50.Ed SOUND DAMPENING DUCT DAMPER

Arthur P. DeLeon, Simi Valley, California
27 July 1999 (Class 251/305); filed 7 March 1997

A plastic or fabric covering over the frame of this spring-loaded damper helps to mute the noise that otherwise would be generated with a sheet metal damper that is creating restricted air flow.—CJR

5,942,736

43.50.Gf ANTINOISE BARRIER WITH TRANSPARENT PANELS, PROVIDED WITH ACOUSTIC INSULATION AND ACOUSTIC ABSORPTION CHARACTERISTICS

Rivo Cortonesi, assignor to Dieselbox SA; Vismara Attillo s.r.l.
24 August 1999 (Class 181/289); filed in Italy 9 July 1997

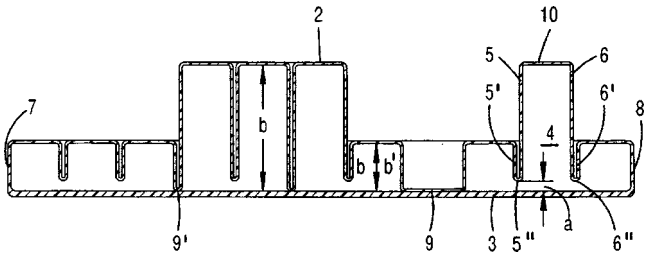
This noise barrier has two parallel panels with an airspace between them. One panel has a perforated face with holes having differing dimensions; the other panel has no holes. The air cylinder between the two panels at each hole of the perforated face panel acts as a Helmholtz acoustic resonator to absorb sound.—CJR

5,959,264

43.50.Gf SOUND ABSORBER

Eduard Bruck and Joachim Oetken, West Germany
28 September 1999 (Class 181/286); filed in Germany 12 August 1994

A folded plate of sheet metal, properly welded together with openings between adjacent cells, presents a common air space enclosed between the structure and the bottom portion, and this in turn helps absorb sound, like a



Helmholtz resonator. This tuned absorber might be used around an engine block.—CJR

5,959,265

43.50.Gf LAMBDA/4-WAVE SOUND ABSORBER

Robert H. Van Ligten, assignor to Rieter Automotive (International) AG
28 September 1999 (Class 181/286); filed in Switzerland 27 January 1995

This absorber panel is designed for vehicles, and comprises a large number of tubular Helmholtz resonators. The array of openings is staggered and distributed so as to provide a range of absorption and structural rigidity.—CJR

5,962,823

43.50.Gf NOISE INSULATING WALL STRUCTURE

Yasuyuki Asahara and Keijiro Iwao, assignors to Nissan Motor Company, Limited
5 October 1999 (Class 181/286); filed in Japan 17 May 1995

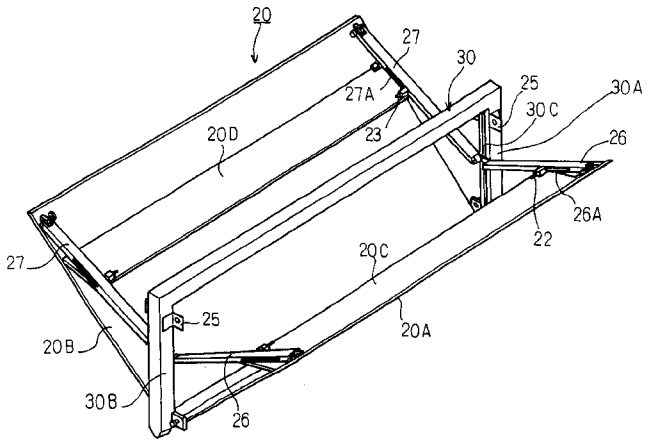
This is a noise insulating wall perhaps used for an under cover of an automobile, that is particularly permeable, and allows dirt and moisture to drain out.—CJR

5,971,096

43.50.Gf NOISE BARRIER AND METHOD OF INSTALLING SAME

Koich Matsumoto *et al.*, assignors to Nihon Doro Kodan; Bridgestone Corporation
26 October 1999 (Class 181/210); filed in Japan 1 December 1997

This noise barrier claims to be able to improve its barrier effectiveness with a V-shaped appendage on top of the barrier, aided by sound absorptive material on the inside surfaces of the appendage. The basic panel is com-



prised of H-shaped support columns that accept panels that stack in the groves of the column.—CJR

5,971,099

43.50.Gf SOUNDPROOF MATERIAL

Zenichi Yasuda *et al.*, assignors to Toyoda Gosei Company Limited
26 October 1999 (Class 181/286); filed in Japan 17 June 1997

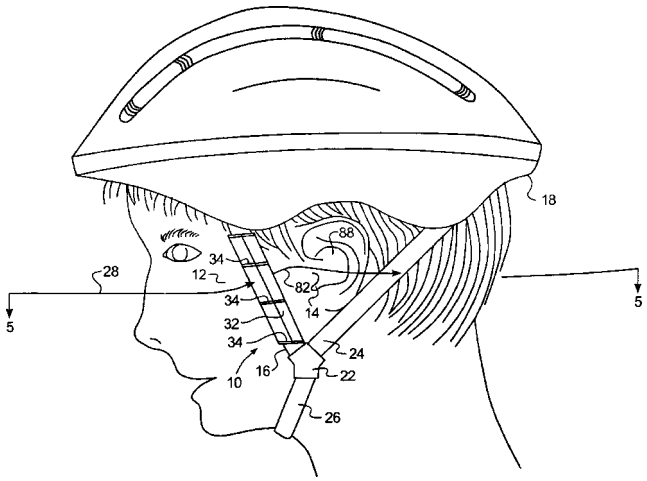
This engine liner uses a substrate and assembly of fibers made from different thermoplastic resins bonded together by ultrasonic vibration to be lightweight and flexible.—CJR

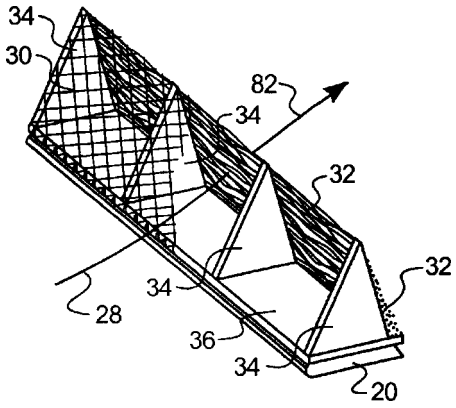
6,029,769

43.50.Gf PERMEABLE AERODAM TO REDUCE WIND NOISE

James B. Tichy, Sausalito, California
29 February 2000 (Class 181/136); filed 18 July 1997

A permeable aerodam (**10**) includes a shaped fiber filter (**32**) designed to impede the motion of air flowing through it. A headwind flows around the side of the head and partially through the aerodam such that an impedance





gradient caused by the shaped filter eliminates the onset of noisy turbulence heard by the rider while allowing ambient sounds to be enjoyed.—JE

6,039,532

43.50.Gf BLOWER FAN BLADE PASSAGE RATE NOISE CONTROL SCHEME

Kenneth G. McConnell, assignor to Iowa State University Research Foundation, Incorporated
21 March 2000 (Class 415/119); filed 17 July 1997

This patent addresses reduction of noise from centrifugal fans or blowers, in which tonal noise results as the air flow through the impeller is restricted abruptly when blades pass the cutoff area, where the impeller circumference comes near the baffle between the impeller and the fan outlet. Noise reduction is achieved by making the flow restriction less abrupt. This may be accomplished by having the lip of the baffle extend at an angle to the axial direction, rather than axially, or by perforating this baffle so that the flow experiences gradually increasing flow restriction.—EEU

5,926,908

43.50.Yw ACOUSTIC COMMUNICATOR FOR CENTRAL VACUUM CLEANERS

Edward W. Lindsay, Jr., assignor to Lindsay Manufacturing, Incorporated
27 July 1999 (Class 15/319); filed 7 June 1995

This device uses a continuous multi-frequency signal, generated by air flow over a reed and transmitted through the pipe system, to allow the operator of a remote intake device to control a central vacuum cleaner. An adaptive detector and signal processor interpret the signal.—CJR

6,041,657

43.50.Yw OUTDOOR NOISE TESTING SYSTEM

Mark B. Sutherland, assignor to Caterpillar, Incorporated
28 March 2000 (Class 73/646); filed 23 December 1997

The system described in this patent is intended for the automatic determination of the total sound power produced by construction machines (such as bulldozers) and related information, in accordance with ISO standards 6393 through 6396. The system consists of a series of microphones that are disposed around an appropriate test volume, plus a processor that accepts the microphone signals and computes the sound power. A position detection arrangement is used to determine when a machine that travels through the test volume is at the appropriate location for sound power data acquisition. Data from an anemometer, a temperature sensor, and a barometer are also recorded. Remote triggering of the data recording process is provided, so that the machine operator can initiate data acquisition when his machine is in the appropriate operating mode.—EEU

5,987,823

43.55.Dy CONTROLLING ACOUSTICS AND EMISSIVITY IN SPORTS ARENAS AND CONCERT HALLS IN WARM CLIMATES IN AIR CONDITIONING MODE

Bernard F. Zarnick, Cleveland, Ohio
23 November 1999 (Class 52/6); filed 1 May 1995

With panels or baffles wrapped in different thicknesses of wrapping and different heat absorbing colors (white or black, for instance) the acoustical and thermal radiating properties of a ceiling can be varied, as for an ice rink in a warm climate.—CJR

5,962,107

43.55.Ev PERFORATED CELLULAR SOUND ABSORPTION MATERIAL

Barry Lynn Lowery *et al.*, assignors to Johns Manville International, Incorporated
5 October 1999 (Class 428/131); filed 29 October 1997

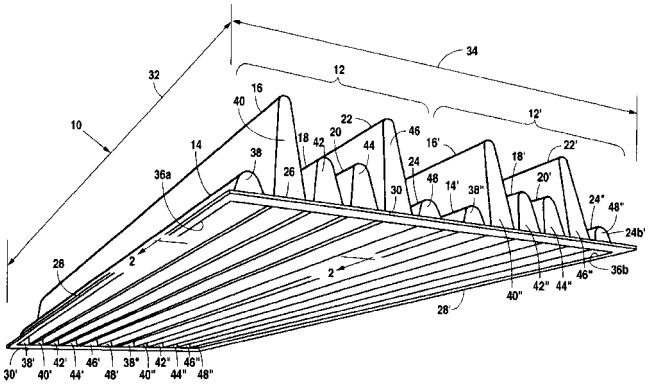
It is claimed that ''needling'' perforations of different depths and diameters and shapes and in different patterns can increase the sound absorptive properties of a low-density closed cell foam. Data shows the increase in normal incidence sound absorption coefficients, mainly in the range below 2000 Hz. Typical foams for which these perforation processes could be beneficial are polyinide, polyamide, and polyethylene closed cell foams of around 2 pounds per cubic foot or less.—CJR

5,969,301

43.55.Ev ACOUSTIC DIFFUSER PANEL SYSTEM AND METHOD

Burton E. Cullum, Jr. and Debbie L. Cullum, Leander, Texas
19 October 1999 (Class 181/286); filed 23 December 1996

This acoustic diffuser panel is an integral unit with a sound reflective surface molded into parabolic curvilinear wells, the size of which is deter-



mined by number theory. An outer lip supports the panel in a suspension system.—CJR

5,929,394

43.55.Ti SOUND ENCLOSURE

John Westerbeke, Jr., assignor to Westerbeke Corporation
27 July 1999 (Class 181/204); filed 7 October 1997

This enclosure creates an air-tight box. The interconnections of the panels are the crux of the design because they allow the top, side, and end panels to be each separately removed without removing the other panels. This enables unrestricted access to all sides of the enclosed equipment, which might typically be a gas or diesel engine and/or generator (perhaps for marine applications).—CJR

6,055,319

43.66.Ts SELECTABLE HANDLE FOR HEARING DEVICES

Adnan Shennib and Dennis A. Oberlander, assignors to Decibel Instruments, Incorporated
25 April 2000 (Class 381/328); filed 6 November 1997

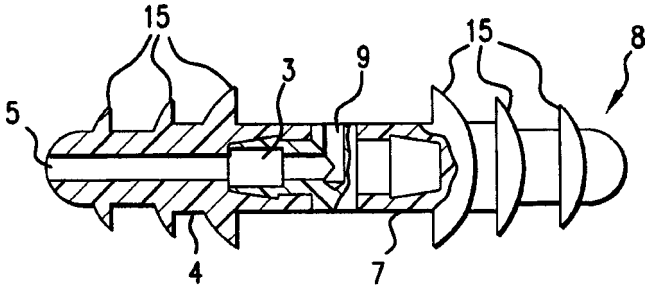
A replaceable removal handle locks into a deeply fitting hearing aid shell, thus eliminating the need for the usual cementing operation to hold the handle in place. The locking mechanism makes it possible for the hearing aid dispenser to utilize handles of different lengths and shapes. Selection of a specific removal handle is based on the cosmetic and dexterity needs of an individual wearer. The handle is replaceable by sliding it sideways into the hearing aid battery compartment.—DAP

5,936,208

43.66.Vt HEARING PROTECTOR AGAINST LOUD NOISES

Pascal Hamery, assignor to Institut Franco-Allemand De Recherches De Saint-Louis
10 August 1999 (Class 181/135); filed in France 18 September 1997

A two-sided insert hearing protector will provide maximum attenuation if one side is inserted and tuned, and lower attenuation if the other side is inserted. Protection from sound up to 190 dB is claimed. Since the skull



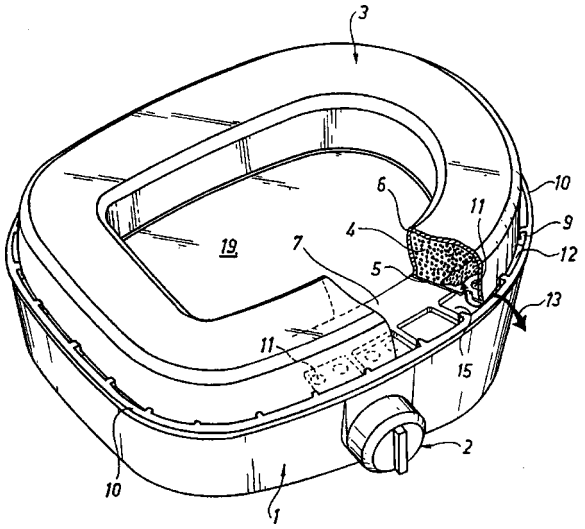
provides maximum attenuation of about 50 dB, the inventor does not explain how this flanking limit might be overcome.—JE

5,970,160

43.66.Vt EARMUFF

Lars Bertil Nilsson and Per Olof Hiselius, assignors to Dalloz Safety AB
19 October 1999 (Class 381/371); filed in Sweden 1 February 1995

This circumaural earmuff hearing protector contains ducts beneath the resilient cushion that can be modified by changing the sealing ring (**15**) on which it is mounted. By changing the size and number of ducts (**11**), it is



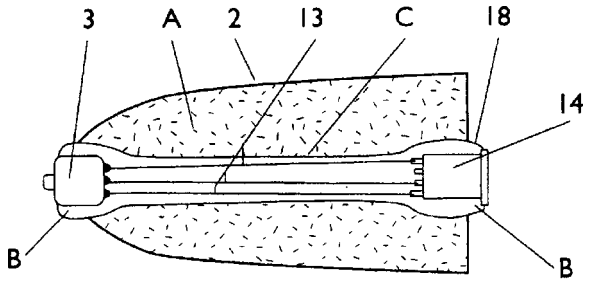
possible to change the noise transmission under the cushion to the interior of the muff and, therefore, the attenuation of the protector.—JE

5,970,795

43.66.Vt APPARATUS AND METHOD FOR TESTING ATTENUATION OF IN-USE INSERT HEARING PROTECTORS

Michael F. Seidmann *et al.*, assignors to Sound Technologies, Incorporated
26 October 1999 (Class 73/585); filed 14 November 1994

This apparatus employs an insert hearing protector with an integral microphone (**3**) mounted adjacent to the eardrum to measure sound pressure in the ear canal. To determine attenuation of the hearing protector, the



incident sound pressure must be determined by other means. An insert is provided to replace the microphone when the measurement has been completed.—JE

5,979,451

43.66.Vt EARMUFF SOUND PROTECTOR

Howard S. Leight, assignor to Howard S. Leight and Associates, Incorporated
9 November 1999 (Class 128/864); filed 9 April 1987

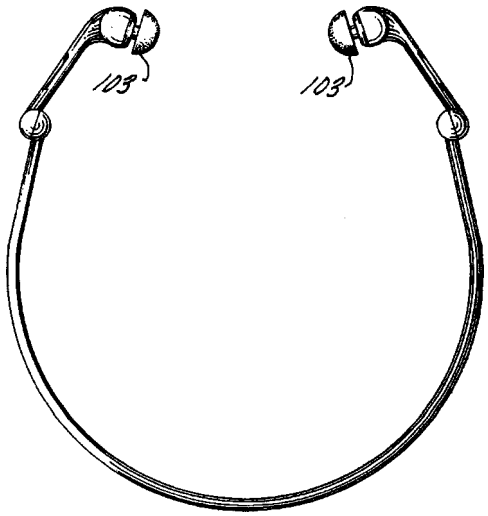
This describes a moldable insert earplug hearing protector and the use of the same material for an earmuff cushion. It is a continuation of an earlier abandoned patent.—JE

5,983,399

43.66.Vt HEARING PROTECTION DEVICE

Robert N. Falco and Marc Doty, assignors to Cabot Safety Intermediate Corporation
16 November 1999 (Class 2/209); filed 15 August 1996

This hearing protection device differs from others by the inclusion of a



multi-degree of freedom articulated support band.—JE

5,996,123

43.66.Vt EARMUFF FOR NOISE BLOCKING

Howard S. Leight and Edwin Woo, assignors to Bacon USA Safety, Incorporated
7 December 1999 (Class 2/209); filed 16 October 1998

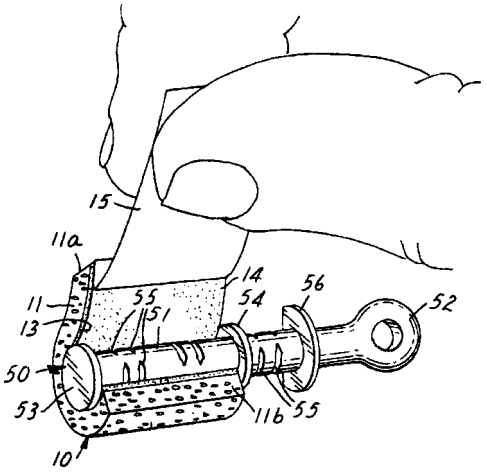
An earmuff-type hearing protector with composite layer internal sound absorption and slow recovery foam cushions claimed to distribute pressure more evenly about the pinna.—JE

5,996,584

43.66.Vt SEALING STRIP FOR EAR PLUGS AND THE LIKE

Robert J. Oliveira and Davis W. Chamberlin, assignors to Hearing Components, Incorporated
7 December 1999 (Class 128/864); filed 5 March 1998

A system for producing a disposable insert hearing protector consist-



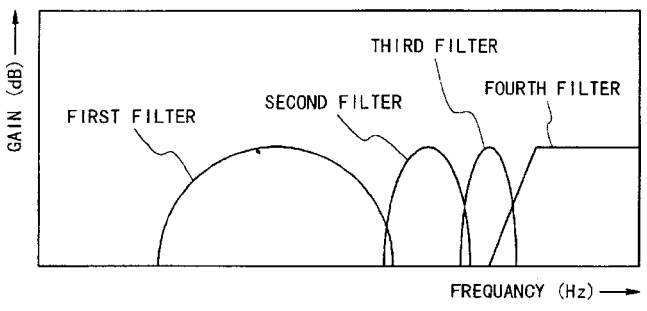
ing of a reusable core (**50**) of different geometries depending on ear canal size and replaceable, compressible foam strips (**11**).—JE

6,035,045

43.71.An SOUND IMAGE LOCALIZATION METHOD AND APPARATUS, DELAY AMOUNT CONTROL APPARATUS, AND SOUND IMAGE CONTROL APPARATUS WITH USING DELAY AMOUNT CONTROL APPARATUS

Akihiro Fujita *et al.*, assignors to Kabushiki Kaisha Kawai Gakki Seisakusho
7 March 2000 (Class 381/17); filed in Japan 22 October 1996

This audio processing system accepts a monaural audio input signal along with a source location signal. First, an interaural time delay is computed based on the specified source direction information. Then the direction signal is used to retrieve gain coefficients for four bands of a head-related



transfer function (HRTF). Shown schematically in the figure, the bands are second-order bandpass and high-pass filters. A lookup table provides amplitude values for the four bands in 10-degree increments.—DLR

6,026,361

43.71.Gv SPEECH INTELLIGIBILITY TESTING SYSTEM

Susan L. Hura, assignor to Lucent Technologies, Incorporated
15 February 2000 (Class 704/270); filed 3 December 1998

This device is an automated laboratory system for generating stimulus pairs to test the intelligibility of a speech output system. In use, a particular vowel or consonant feature contrast is specified. The system then generates a list of word pairs, along with the setup required to play those pairs to a test subject.—DLR

6,035,270

43.71.Gv TRAINED ARTIFICIAL NEURAL NETWORKS USING AN IMPERFECT VOCAL TRACT MODEL FOR ASSESSMENT OF SPEECH SIGNAL QUALITY

Michael P. Hollier *et al.*, assignors to British Telecommunications public limited company
7 March 2000 (Class 704/202); filed in European Patent Office 27 July 1995

This device is a test set designed to test telecommunications pathways which include perception-based speech compression. For various reasons, the test signal is to be speechlike, rather than arbitrary tones. If a real speech signal is used, little is learned of the condition of the system under test. The solution is to distort a speech signal with unnatural distortion components. Both the original speech and the distorted version are transmitted and compared at the receiver. The differences are processed by a neural network to characterize the line condition.—DLR

6,026,356

43.72.Ar METHODS AND DEVICES FOR NOISE CONDITIONING SIGNALS REPRESENTATIVE OF AUDIO INFORMATION IN COMPRESSED AND DIGITIZED FORM

H. S. P. Yue *et al.*, assignors to Nortel Networks Corporation
15 February 2000 (Class 704/201); filed 3 July 1997

Nonspeech sounds do not fare very well in most speech coder systems. This system for low bitrate communication detects whether a frame contains speech or other sounds. Speech is handled with a typical vocoder arrangement, while other sounds are processed by a system which takes a longer history into consideration, either for the linear prediction coefficients or by averaging multiple LP coefficient sets. The method is intended primarily for cell phone use, to improve the transmitted background sounds.—DLR

6,026,357

43.72.Ar FIRST FORMANT LOCATION DETERMINATION AND REMOVAL FROM SPEECH CORRELATION INFORMATION FOR PITCH DETECTION

Mark A. Ireton and John G. Bartkowiak, assignors to Advanced Micro Devices, Incorporated
15 February 2000 (Class 704/207); filed 15 May 1996

This pitch detection device measures the dominant formant energy, typically the first formant, as a way of improving the measurement of the voice pitch. A second-order linear prediction analysis provides a formant estimate. The frequency value is converted to a period measure, which is applied as a correction to the autocorrelation function. Any remaining autocorrelation peak is more likely to represent the voice pitch.—DLR

6,029,129

43.72.Ar QUANTIZING AUDIO DATA USING AMPLITUDE HISTOGRAM

Scott A. Kliger *et al.*, assignors to Narrative Communications Corporation
22 February 2000 (Class 704/230); filed 24 May 1996

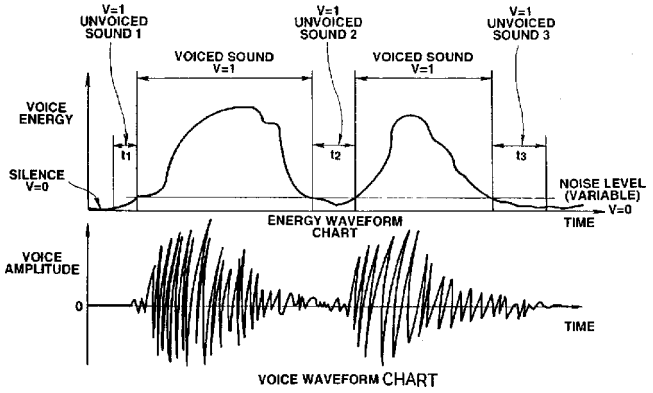
This method of speech compression uses an adaptive system of requantization of the sampled data in which the quantization levels are based on a histogram evaluation. This provides a better statistical estimate of the appropriate quantization levels, reducing the number of bits required for an adequate encoding.—DLR

6,031,915

43.72.Ar VOICE START RECORDING APPARATUS

Hideo Okano *et al.*, assignors to Olympus Optical Company, Limited
29 February 2000 (Class 381/56); filed in Japan 19 July 1995

This voice-activated switch is designed to minimize the chance of false triggering due to input noise. An analog voice input circuit generates a voltage representing the energy level. At the same time, the input signal is analyzed by a DSP chip and compressed for storage in a memory device. If



the energy level value exceeds a threshold value for a determined number of analysis frames, then the compressed data is transferred to the memory unit.—DLR

6,032,115

43.72.Ar APPARATUS AND METHOD FOR CORRECTING THE DIFFERENCE IN FREQUENCY CHARACTERISTICS BETWEEN MICROPHONES FOR ANALYZING SPEECH AND FOR CREATING A RECOGNITION DICTIONARY

Hiroshi Kanazawa *et al.*, assignors to Kabushiki Kaisha Toshiba
29 February 2000 (Class 704/234); filed in Japan 30 September 1996

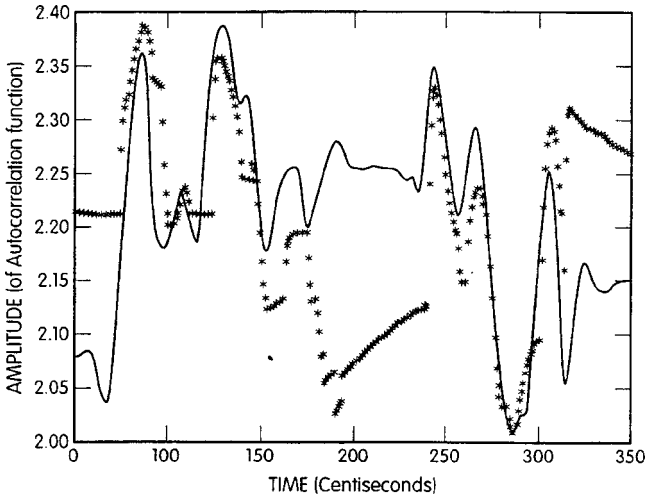
A microphone embedded in the cover or keyboard of a laptop computer is generally not well situated for proper voice pickup. This patent describes a procedure by which such a mic can be calibrated and filtered so as to provide a corrected speech signal to a speech processor in the laptop. A standard reference microphone is placed near the laptop in such a way as to simultaneously record a signal with the laptop mic. The known characteristics of both the standard mic and the mic which was originally used to record the recognition system training data are used to construct a correction filter for the laptop mic.—DLR

6,035,271

43.72.Ar STATISTICAL METHODS AND APPARATUS FOR PITCH EXTRACTION IN SPEECH RECOGNITION, SYNTHESIS AND REGENERATION

ChengJun Julian Chen, assignor to International Business Machines Corporation
7 March 2000 (Class 704/207); filed 15 March 1995

This pitch analyzer computes a normalized autocorrelation function, and then, in each frame, marks three candidate peaks for further analysis. Based on a theory that the vocal cords cannot change their effective resonance condition rapidly, whether or not voicing is currently in progress, a



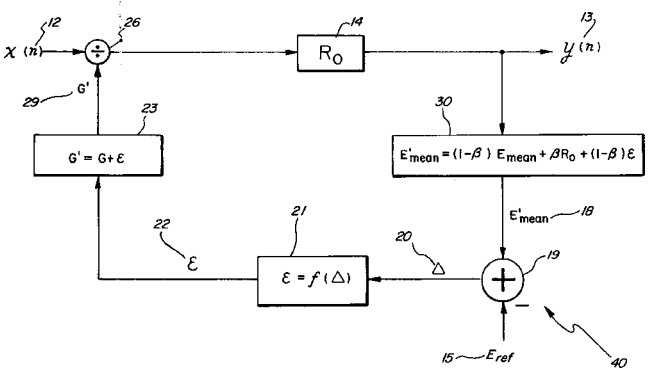
theoretically continuous pitch curve is taken as a tracking target. Using a system of anchor points and forward and backward running mean and average pitch values, a pitch curve is found which best fits the continuous target curve.—DLR

6,037,834

43.72.Ar METHOD AND APPARATUS FOR AUTOMATIC GAIN CONTROL WITH IMPROVED RESPONSE TIME AND STABILITY

Bhasker P. Patel and Kenneth E. Garey, assignors to Conexant Systems, Incorporated
14 March 2000 (Class 327/560); filed 6 March 1998

You thought all possible gain control circuits had already been invented. Maybe you were right. In this circuit, the difference calculation includes an extra term which depends on the result of that calculation. Since the figure does not show where the delay element is located, we can only



guess how it really works. The authors assert that the new change is incorporated into the error signal, allowing a more rapid response to an input level change.—DLR

6,026,358

43.72.Bs NEURAL NETWORK, A METHOD OF LEARNING OF A NEURAL NETWORK AND PHONEME RECOGNITION APPARATUS UTILIZING A NEURAL NETWORK

Hideto Tomabechi, assignor to Justsystem Corporation
15 February 2000 (Class 704/232); filed in Japan 22 December 1994

This patent describes several related arrangements for using a neural network (NN) to recognize phonemes in an input speech signal. The two major additions to a typical NN architecture are a feedback or context layer connected to the input and a hypothesis vector layer connected to the output. These two extra layers do not have weights, but represent preset ''value'' or bias vectors. The context layer is set from the previous state hidden layer values. During training, the hypothesis layer is preset with a binary code representing the target phoneme.—DLR

6,038,531

43.72.Bs SIMILAR WORD DISCRIMINATION METHOD AND ITS APPARATUS

Yasunaga Miyazawa *et al.*, assignors to Seiko Epson Corporation
14 March 2000 (Class 704/232); filed in Japan 30 January 1997

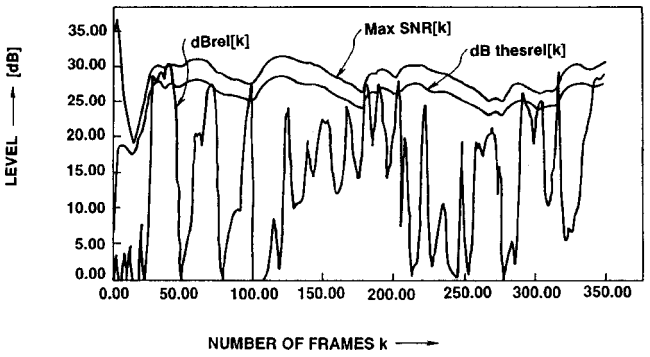
This word spotter speech recognition system uses a dynamic recurrent neural network as the primary word spotting system. The neural net architecture is not further described. In addition to the neural net, the system uses a codebook based on the tenth-order cepstrum acoustic data to further distinguish similar input words. If the neural net score for a potentially confusable input word is above a preset threshold, then the codebook is checked to provide a more reliable word identification.—DLR

6,032,114

43.72.Dv METHOD AND APPARATUS FOR NOISE REDUCTION BY FILTERING BASED ON A MAXIMUM SIGNAL-TO-NOISE RATIO AND AN ESTIMATED NOISE LEVEL

Joseph Chan, assignor to Sony Corporation
29 February 2000 (Class 704/226); filed in Japan 17 February 1995

This noise reduction system is a variation on an old and well-known method of signal versus noise level evaluation, but adds several steps to the procedure. A noise spectrum is estimated by known methods and the ratio of the overall signal to this noise estimate is computed. A frame-maximum



signal-to-noise ratio is determined for each frame and this value is used to control the application of a filter derived from the signal-to-noise spectral ratio.—DLR

6,035,048

43.72.Dv METHOD AND APPARATUS FOR REDUCING NOISE IN SPEECH AND AUDIO SIGNALS

Eric John Diethorn, assignor to Lucent Technologies, Incorporated
7 March 2000 (Class 381/94.3); filed 18 June 1997

This speech signal noise reduction system divides the signal into multiple subbands and processes each subband through a series of steps. Essentially, the subband operations effect a change of band gain depending on the relative speech and noise levels. Initial level estimates are obtained using typical, known methods. Signal levels are then reevaluated using bandwidths slightly greater or less than the subband, all with various attack and decay rates. The various level measures are then combined to produce a final subband adjustment before the bands are recombined to produce the output signal.—DLR

6,029,124

43.72.Fx SEQUENTIAL, NONPARAMETRIC SPEECH RECOGNITION AND SPEAKER IDENTIFICATION

Laurence S. Gillick *et al.*, assignors to Dragon Systems, Incorporated
22 February 2000 (Class 704/200); filed 21 February 1997

This patent covers a system to collect and process training speech data such that both speech and speaker information can be recovered from a new test speech sample. Both training and test acoustic data are converted to cepstral first and second differences. The essence of the system is that instead of reducing the training data to a particular speech model, such as a hidden Markov model, the training data itself is maintained and compared directly with the incoming speech. This allows more accurate large-vocabulary recognition or speaker identification. It is said that the system can also perform language identification.—DLR

6,026,360

43.72.Gy SPEECH TRANSMISSION/RECEPTION SYSTEM IN WHICH ERROR DATA ARE REPLACED BY SPEECH SYNTHESIZED DATA

Hiroshi Ono, assignor to NEC Corporation
15 February 2000 (Class 704/260); filed in Japan 28 March 1997

Primarily intended for a situation in which a retransmission request is not possible, this speech coder is made more robust by transmitting both the normal CELP-coded compressed speech data along with a text representation suitable for resynthesis. If error detection at the receiver indicates that the compressed data have been corrupted and the text data are still valid, then a synthetic voice is generated as a backup version to maintain communications.—DLR

6,035,273

43.72.Gy SPEAKER-SPECIFIC SPEECH-TO-TEXT/ TEXT-TO-SPEECH COMMUNICATION SYSTEM WITH HYPERTEXT-INDICATED SPEECH PARAMETER CHANGES

Carl Joseph Spies, assignor to Lucent Technologies, Incorporated
7 March 2000 (Class 704/270); filed 26 June 1996

This patent describes a text vocoder, a device with a long history, but only recently achieving usable quality levels. In this version, the transmitter contains a permanent memory of the talker's voice characteristics. Upon establishing a connection, speaker defining packets are interchanged, allowing the receiver to alter the output to closely match the speaker's voice. The transmitted text is augmented by hypertext characters which indicate current values for pitch, speech rate, and loudness. A transmission rate of 160 to 300 bits per second is estimated.—DLR

6,029,131

43.72.Ja POST PROCESSING TIMING OF RHYTHM IN SYNTHETIC SPEECH

Edward A. Bruckert, assignor to Digital Equipment Corporation
22 February 2000 (Class 704/260); filed 28 June 1996

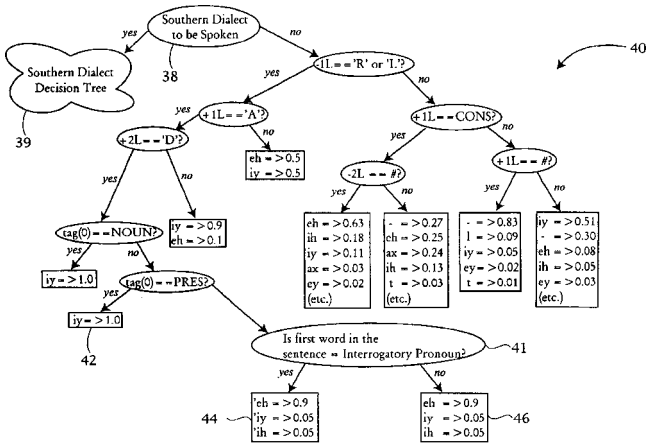
This patent describes a timing generator for producing a more natural-sounding synthetic speech signal. Initially, each phoneme to be synthesized is assigned an intrinsic duration value. A ''natural timing interval'' (NTI) is then established, which may be a syllable duration, the interval between stress points, or another such unit, depending on the timing structure of the synthesized language. The individual phoneme durations are then adjusted according to phoneme-specific compliance values such that the NTI is closely matched.—DLR

6,029,132

43.72.Ja METHOD FOR LETTER-TO-SOUND IN TEXT-TO-SPEECH SYNTHESIS

Roland Kuhn and Jean-Claude Junqua, assignors to Matsushita Electric Industrial Company
22 February 2000 (Class 704/260); filed 30 April 1998

This speech synthesizer system uses two stages of decision tree analysis to examine the spelling of each word, the word class (part-of-speech), and the local syntactic structure to determine the most probable pronunciation for each word. Word-level stress values are also assigned by the decision trees. Different decision trees may be selected by the user in order to produce pronunciations according to a selected dialect.—DLR

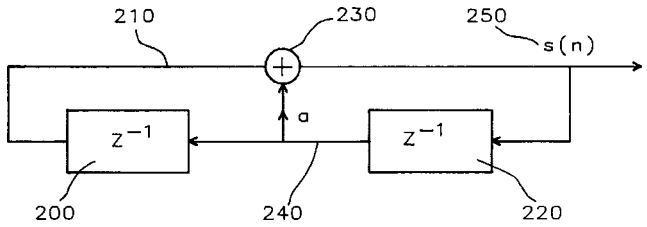


6,029,133

43.72.Ja PITCH SYNCHRONIZED SINUSOIDAL SYNTHESIZER

Ma Wei, assignor to Tritech Microelectronics, Limited
22 February 2000 (Class 704/265); filed 15 September 1997

This speech signal generator is intended for use as a vocoder receiver, in which speech parameter frames are received and used to produce speech output. Amplitudes for each harmonic are computed from the spectral data and a frame-related phase value is determined. Each harmonic is also assigned a voicing value, which, if unvoiced, presumably causes a phase randomization for the frame. A bank of sine-wave generators, one of which is shown in the figure, produces a sinusoid for each harmonic. These are summed to produce the output signal.—DLR

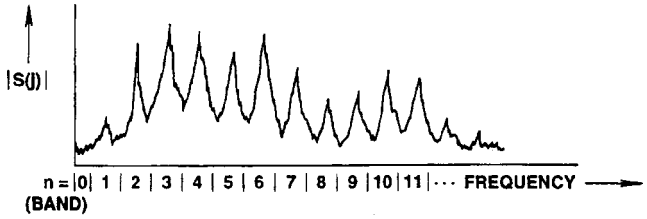


6,029,134

43.72.Ja METHOD AND APPARATUS FOR SYNTHESIZING SPEECH

Masayuki Nishiguchi and Jun Matsumoto, assignors to Sony Corporation
22 February 2000 (Class 704/268); filed in Japan 28 September 1995

This speech signal generator is intended for use with a vocoder, which would supply a sequence of speech data frames for synthesis. Each frame would include a pitch value and a voiced/unvoiced indicator. An amplitude value for each harmonic is determined from the spectral envelope. For a voiced frame, harmonic phases are computed so as to be continuous across frames. For an unvoiced frame, the phases are set at predetermined values.—DLR



6,033,224

43.72.Ja READING MACHINE SYSTEM FOR THE BLIND HAVING A DICTIONARY

Raymond C. Kurzweil *et al.*, assignors to Kurzweil Educational Systems
7 March 2000 (Class 434/112); filed 27 June 1997

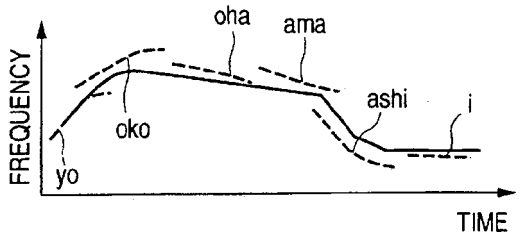
This reading machine is an updated version of a similar machine sold by a company started by the first author two decades ago. The current machine has more controllable modes and, in particular, includes a built-in dictionary from which word definitions may be played back via speech synthesis. Organized for convenient use by a blind operator, the definition may be requested for a word just read by the device.—DLR

6,035,272

43.72.Ja METHOD AND APPARATUS FOR SYNTHESIZING SPEECH

Hirofumi Nishimura *et al.*, assignors to Matsushita Electric Industrial Company, Limited
7 March 2000 (Class 704/258); filed in Japan 25 July 1996

This waveform-segment-concatention speech synthesizer stores triphone units of recorded Japanese speech, each with two pitch accent levels. Each stored unit is marked with the points at which it would best be cut to join up with a neighbor and also includes a phonetic transcription for the unit. As stored units are joined to produce speech, there is no further modification of the output pitch.—DLR

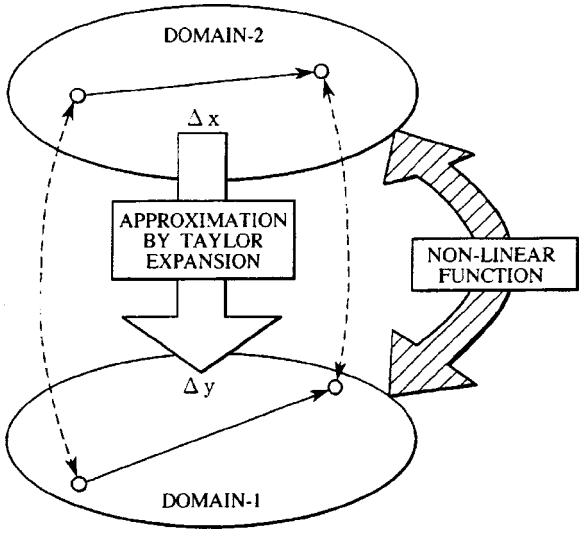


6,026,359

43.72.Ne SCHEME FOR MODEL ADAPTATION IN PATTERN RECOGNITION BASED ON TAYLOR EXPANSION

Yoshikazu Yamaguchi *et al.*, assignors to Nippon Telegraph and Telephone Corporation
15 February 2000 (Class 704/256); filed in Japan 20 September 1996

The hidden Markov model (HMM) method widely used for speech recognition suffers from a degradation of performance when the background noise conditions are mismatched between the training environment and the recognition environment. This patent presents a method of adapting the HMM parameters to adjust a particular training data set to a different noise background. A Taylor series expansion approximates the change in cepstral vectors of the background noise. From this, a Jacobian matrix is computed such that the HMM parameters may be quickly adapted to a new noise background at the time of recognition.—DLR

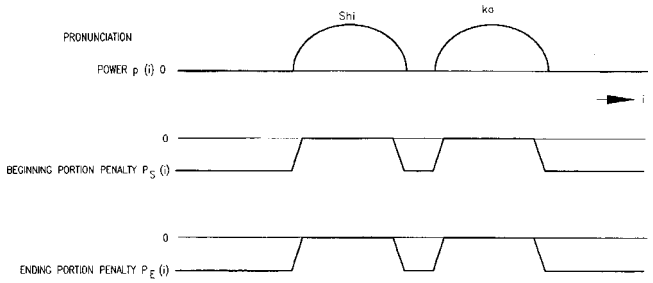


6,029,130

43.72.Ne INTEGRATED ENDPOINT DETECTION FOR IMPROVED SPEECH RECOGNITION METHOD AND SYSTEM

Takashi Ariyoshi, assignor to Ricoh Company, Limited
22 February 2000 (Class 704/248); filed in Japan 20 August 1996

This isolated word recognizer seems to be a fairly traditional dynamic programming template comparison system. An energy measure is computed for each frame and tracked along with the mel cepstral acoustic data. Frame



similarity measures for both energy and spectral data are used together to determine the word boundaries.—DLR

6,032,116

43.72.Ne DISTANCE MEASURE IN A SPEECH RECOGNITION SYSTEM FOR SPEECH RECOGNITION USING FREQUENCY SHIFTING FACTORS TO COMPENSATE FOR INPUT SIGNAL FREQUENCY SHIFTS

Safdar M. Asghar and Lin Cong, assignors to Advanced Micro Devices, Incorporated
29 February 2000 (Class 704/238); filed 27 June 1997

This isolated word recognizer uses a fairly typical arrangement of template matching computations, except for the use of a novel spectral distance measure. The Fuzzy Matrix Quantizer method seems to be based on a weighted sum of the differences in line spectral pairs coefficients. The system is described more fully in the Ph.D. thesis of one of the authors. Spectral differences are computed for each word in the training set. The fuzzy quantizer is a way of computing centroids of the training data, which are in turn used to build codebooks for recognition.—DLR

6,035,017

43.72.Ne BACKGROUND SPEECH RECOGNITION FOR VOICE MESSAGING APPLICATIONS

Francis Michael Fenton *et al.*, assignors to Lucent Technologies, Incorporated
7 March 2000 (Class 379/88.04); filed 24 January 1997

This is an automated telephone response system which, in addition to recording the incoming message, also performs a keyword recognition task. If one or more of certain prestored keywords are seen in the caller's message, the system may initiate a corresponding action sequence. Such action sequences typically consist of a prompt to be played, forwarding the call, sending a paging signal, or activating a local display.—DLR

5,868,683

43.80.Qf TECHNIQUES FOR PREDICTING READING DEFICIT BASED ON ACOUSTICAL MEASUREMENTS

Athanassios Protopapas *et al.*, assignors to Scientific Learning Corporation
9 February 1999 (Class 600/559); filed 24 October 1997

The patent discloses a method to predict reading deficits based on the ability to process phonemes. It assumes that difficulties in speech processing will lead to difficulties in reading. The phoneme discrimination test is based on the ability to discriminate changes in frequency and level of test stimuli.—JE

Transient response of quasi-isotropic fiber composite laminates to internal line sources

E. Rhian Green
Department of Engineering, The University of Leicester, Leicester LE1 7RH, United Kingdom

W. A. Green
Department of Mathematical Sciences, Loughborough University, Loughborough, Leicester LE11 3TU, United Kingdom

(Received 25 February 2000; accepted for publication 7 August 2000)

This article considers the propagation of elastic waves in an eight-ply quasi-isotropic laminate arising from line sources of dislocation located at each of the seven interfaces in turn. The laminate is composed of identical layers of a fiber composite material which is modeled as a homogeneous transversely isotropic elastic continuum with the axis of transverse isotropy along the fiber direction. The line source sets up a straight crested wave traveling along the laminate in the direction normal to the load line and the elastodynamic equations within each layer are solved by taking the Laplace transform with respect to time and the Fourier transform with respect to the spatial coordinate in the direction of propagation. The resulting system of six first-order differential equations in each layer is solved to obtain the transforms of the displacement and stress components throughout the laminate. The time history of any displacement or stress component at any location may then be recovered by numerical inversion of the double transform. The graphs presented show the time history of the normal displacement of the top surface of the laminate at distances of 1 and 20 plate thicknesses from the plane of action of the sources. These graphs are for four different orientations of the line of action of the sources, namely, at angles 0, 30, 60, and 90 degrees to the fiber direction in the surface layer. The numerical inversion involves a summation over different modes of Rayleigh–Lamb waves in the laminate and results are also presented showing the contributions to the overall response from some of the individual modes.
[S0001-4966(00)03011-3]

PACS numbers: 43.20.Mv [DEC]

I. INTRODUCTION

In fiber composite laminated materials the acoustic emission technique has the potential for not only locating the source of the emission but also for determining the nature of the fracture event which gave rise to the received signals. The pulses arising from events such as fiber breakage, matrix cracking, fiber/matrix debonding and ply delamination are completely different in character and give rise to different signals at the receivers. By recording the time history of these signals it should be possible to determine the nature of the initiating event. The performance of this task involves the solution of the inverse problem of elastodynamics in a multi-layered, anisotropic inhomogeneous medium—a problem which as yet remains to be solved. An alternative approach, which avoids solving the difficult inverse problem, is to attempt to replicate the receiver signal by simulating the time history of sources of known characteristics and modifying the parameters of the source function in order to achieve good agreement with observed waveforms. This approach has been adopted in Ref. 1 for isotropic materials and in Ref. 2 for fiber composite materials and it has been shown to be successful when applied to a number of different fracture events. In the work reported in Ref. 1, the source characteristics and the relative positions of source and receiver are taken as known, and the simulated signals are based on the earliest arrivals at the receiver, for which the Green's function is that associated with the surface response in a half-space of isotropic elastic material. In order to simulate the receiver signal in the case of the fiber composite laminate it is necessary to determine the Green's functions associated with a range of initiating sources, located at various levels within the structure. The simulation method adopted in Ref. 2 employs classical laminate theory with shear correction in order to predict the response to internal point loads in terms of the Rayleigh–Lamb modes of wave propagation. It is pointed out in Ref. 2 that laminate theory yields the response arising from the fundamental Rayleigh–Lamb mode only and the results are accordingly only valid for low frequencies. To obtain the simulated signals for higher frequencies it is necessary to solve the full elastodynamic equations in each of the plies and this is done here using a multiple transform method. The numerical results which are presented involve a summation over 20 different modes of the Rayleigh–Lamb waves in the laminate, and graphs are also included to illustrate the contribution to the overall response from some of the individual modes.

The study of wave propagation in fibrous laminates involves two main differences from that in monolithic isotropic materials. First, the presence of long parallel fibers embedded in an isotropic matrix make the basic composite material both inhomogeneous and anisotropic, so that the waves undergo scattering and attenuation and their speeds vary with the direction of propagation. Second, the presence

of the interfaces between the layers of the laminate give rise to multiple reflections and lead to a complex pattern of waves arriving at the surface. The problem of inhomogeneity can be overcome to a large extent by restricting consideration to disturbances whose wavelengths are large compared with a typical fiber diameter. In this circumstance, the scattering effect of the fibers will be small and the composite material forming each layer of the laminate may be modeled as a homogeneous, but anisotropic continuum. Typically, the fibers in a carbon fiber/epoxy resin prepreg have diameters of the order of 6 μm while the prepreg has depth of the order of 125 μm. Disturbances with wavelengths of the order of some 60 μm or greater, corresponding to frequencies of order 25 MHz or less, will undergo little scattering from the fibers while still being small enough to be affected by the depth of the prepreg and to undergo reflection at the interfaces.

While the restriction on wavelengths brings the simplification of homogeneity, the anisotropic nature of the composite makes the problem of elastic wave motion significantly more difficult than for isotropic materials. Thus, the disturbance emanating from a point source is not radially symmetric and the signal is a function of all three space variables as well as the time. The determination of the Green's functions arising from point sources is consequently a computationally demanding task and to date this has been carried out for surface loadings for which the spectra are negligible above some 5 MHz.[3,4] A less demanding approach to understanding the effects of anisotropy and which covers an extended frequency range can be obtained by examining the transmission of disturbances arising from line loads within or on the surface of the laminate.[5,6] Line loads of uniform strength will set up a motion in the form of straight crested waves, traveling in the plane of the laminate with speeds which vary with the orientation of the line relative to the preferred material directions. In this wave motion the spatial dimension of the problem is reduced by one, with the stress and displacement being functions of time and two space coordinates. This simulation method involves expressing the transient solution in terms of the modes of propagation of guided time harmonic (Rayleigh–Lamb) waves in the laminate. There is an extensive literature dealing with these propagation modes and a list of references is contained in the review article by Chimenti.[7]

This paper adopts the transversely isotropic continuum model in each ply, in order to consider the propagation of stress waves in a quasi-isotropic laminate consisting of eight layers of a fiber composite material. Using transform techniques, closed form solutions are derived for the surface displacements resulting from internal line sources located at one of the interfaces between the plies and making different angles with the fiber directions in the outer layer. The source has the form of a step function in time and consists of a line discontinuity in the component of displacement normal to the plane of the interface. The surface response as a function of time at varying distances from the source is calculated by numerical inversion of the transform solutions. Results are shown for sources located at each of the seven interfaces between the plies and at receiver distances of 1 and 20 laminate thicknesses from the plane of action of the source. It is shown that, in some circumstances, it is possible for the main contribution to the received signal to arise from one of the higher modes of the Rayleigh–Lamb waves.

The two-dimensional solutions arising from line loads provide detailed information about the effects of orientation on the nature of the disturbance. Nevertheless, the problem of practical significance remains the evaluation of the response to an internal point source. One possible approach to determining this Green's function numerically is to integrate the response to a line load over all orientations of the line and preliminary work on developing this approach has been carried out for some simpler model problems.[8,9] Thus, the work reported here, in addition to providing detailed information about the variation of the surface response with source orientation, also constitutes a step towards the evaluation of the point source response.

This work relates to a perfectly elastic model of the composite material and no account has been taken of any damping in the model. Since damping is in general frequency dependent, it will lead to a reduction in both the amplitudes and frequency content of the transmitted signals with increasing distance of travel (see Ref. 10), and the simulations shown here do not include these dissipative effects. The sharp step function input source which is employed in the simulation introduces the complete range of frequencies from zero to infinity. It may well be that the frequency content of acoustic emission signals is rarely above some 5 MHz. In this case an approach such as that of Lih and Mal,[3] based on a modified classical laminate model with damping, may serve to simulate the transients arising from these events. Nevertheless, for a complete understanding of the nature of wave propagation in anisotropic laminated plates it is essential to employ a model which accounts for the presence of the individual laminae and the work presented here forms a step towards the achievement of this goal.

II. THEORY

The plate is constructed from eight layers each of depth h and of infinite lateral extent, of an unidirectional fiber reinforced material modeled as a transversely isotropic elastic continuum, with the axis of transverse isotropy being parallel to the fiber direction. A Cartesian coordinate system of axes $Ox_1x_2x_3$ is chosen, with the origin O in the middle surface of the laminate, Ox_1 in the direction of the upward normal to the plate and Ox_3 along the reference axis, which is taken to be along the fiber direction in the outer ply. The plies are arranged in the quasi-isotropic lay-up $(0 \text{ degrees}, 45 \text{ degrees}, -45 \text{ degrees}, 90 \text{ degrees})_s$ and a line source of uniform strength acts at the interface between two of the plies with the line of action being oriented at an angle $\pi/2-\gamma$ relative to the reference axis. This sets up a straight crested disturbance traveling an angle $-\gamma$ relative to Ox_3 and it follows that the components of displacement and stress are functions of the time t, of the normal coordinate x_1 and of the in-plane spatial variable $x=x_2 \sin\gamma+x_3\cos\gamma$.

The method adopted in order to solve the transient elastodynamic problem is due to Mal[11,12] and it involves taking the Laplace transform with respect to t and the Fourier transform with respect to the spatial variable x in the form

$$U_i(x_1,k,s)=\int_{-\infty}^{\infty}\int_0^{\infty}u_i(x_1,x,t)e^{-st}e^{-ikx}\,dt\,dx,$$

$$(i=1,2,3), \tag{1}$$

$$T_{ij}(x_1,k,s)=\int_{-\infty}^{\infty}\int_0^{\infty}t_{ij}(x_1,x,t)e^{-st}e^{-ikx}\,dt\,dx,$$

$$(i,j=1,2,3). \tag{2}$$

The equations of motion within any layer m are reduced to a system of six first-order differential equations which have the form

$$\frac{d\mathbf{Z}_m}{dx_1}=\mathbf{A}_m\mathbf{Z}_m, \tag{3}$$

where $\mathbf{A}_m$ is a known matrix whose elements depend on the elastic constants, the density, the angle γ, the orientation of the ply and the transform parameters k and s. In Eq. (3) the vector $\mathbf{Z}_m$ is given by

$$\mathbf{Z}_m=(U_1\,U_2\,U_3\,T_{11}\,T_{12}\,T_{13})^T, \tag{4}$$

and ^{T} denotes the transpose of the vector. The solution within each layer is given in terms of six arbitrary constants in the form

$$\mathbf{Z}_m(x_1,k,s)=\mathbf{P}_m\mathbf{E}_m(x_1)\mathbf{K}_m, \tag{5}$$

where $\mathbf{P}_m$ is the matrix whose columns are the eigenvectors of $\mathbf{A}_m$, $\mathbf{K}_m$ is the vector of arbitrary constants and the diagonal matrix $\mathbf{E}_m(x_1)$ has the form

$$\mathbf{E}_m(x_1)=\mathrm{diag}\{e^{kp_1^m(x_1-h_m)},e^{kp_2^m(x_1-h_m)},e^{kp_3^m(x_1-h_m)},$$
$$e^{kp_1^m(h_{m-1}-x_1)},e^{kp_2^m(h_{m-1}-x_1)},e^{kp_3^m(h_{m-1}-x_1)}\}. \tag{6}$$

Here $\pm kp_1^m$, $\pm kp_2^m$, $\pm kp_3^m$ are the six eigenvalues of $\mathbf{A}_m$ and h_{m-1}, h_m are the values of x_1 at the bottom and at the top of layer m, respectively.

The continuity conditions at the interface between layer r and layer $r+1$ are given by

$$\mathbf{P}_{r+1}\mathbf{E}_{r+1}(h)\mathbf{K}_{r+1}=\mathbf{P}_r\mathbf{E}_r(h)\mathbf{K}_r+\mathbf{F}_r, \tag{7}$$

where $\mathbf{F}_r$ denotes the discontinuity (if any) in the vector $\mathbf{Z}$ across the interface. For perfect bonding across the interface the vector $\mathbf{Z}$ is continuous and $\mathbf{F}_r$ is zero. Nonzero components of $\mathbf{F}_r$ correspond to sources acting at the interface. Thus, a nonzero value of the first component represents a crack opening dislocation source while a nonzero value of the fourth component represents a vertical line load acting at the interface. Imposing the conditions (7) across each of the interfaces, together with the traction-free boundary conditions at the top and bottom surfaces of the laminate, leads to a system of 48 linear simultaneous equations for the 48 arbitrary constants. This system of equations has the form

$$\mathbf{M}(k,s)\mathbf{K}=\mathbf{F}G(s), \tag{8}$$

where $\mathbf{M}(k,s)$ is a square matrix of order 48, $\mathbf{K}$ is the column vector of arbitrary constants and $\mathbf{F}$ is a column vector whose components arise from the discontinuities acting across the interfaces. The matrix $\mathbf{M}$ in Eq. (8) is a function of the elastic constants, the material density, the angle of propagation γ, the ply orientations and the transform parameters, and $G(s)$ is the Laplace transform of the time history of the source. The solution of Eqs. (8) may be written in the form

$$\mathbf{K}=\mathbf{M}^{-1}(k,s)\mathbf{F}G(s)=\frac{\mathbf{adj}\,\mathbf{M}(k,s)}{\det\mathbf{M}(k,s)}\mathbf{F}G(s), \tag{9}$$

where $\mathbf{adj}\,\mathbf{M}$ is the adjoint matrix of $\mathbf{M}$ and $\det\mathbf{M}$ is the determinant of $\mathbf{M}$.

The transforms of the displacement and stress components at any location in the laminate are expressed as linear combinations of the elements of the vector $\mathbf{K}$. In particular, the transform $U_1(4h,k,s)$ of the normal displacement at the upper surface $x_1=4h$ has the form

$$U_1(4h,k,s)=\frac{N(k,s)}{\det\mathbf{M}(k,s)}G(s), \tag{10}$$

where $N(k,s)$ is a known function of the transform parameters. The formal solution for the upper surface normal displacement is given by the inverse of the double transform,

$$u_1(4h,x,t)=\frac{1}{4\pi^2 i}\int_{-\infty}^{\infty}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{N(k,s)}{\det\mathbf{M}(k,s)}G(s)e^{st}e^{ikx}\,ds\,dk. \tag{11}$$

The integral with respect to s can be evaluated in terms of the residues of the integrand at the zeros of the function $\det\mathbf{M}(k,s)$ to the left of the line $s=\gamma\,(\mathrm{real}\,\gamma>0)$ in the complex s plane. These zeros all occur on the imaginary axis $s=i\omega$. The equation,

$$\det\mathbf{M}(k,i\omega)=0, \tag{12}$$

is the generalized Rayleigh–Lamb dispersion equation relating the angular frequency ω to the wave number k for plane harmonic waves propagating in the laminate. Note that the wave number k is confined to real values. Equation (12) has an infinite number of pairs of roots $\omega=\pm\omega_j(k)$, $j=1,2,\ldots$. Each curve $\omega=\omega_j(k)$ corresponds to one branch of the dispersion equation and in terms of these roots Eq. (11) becomes

$$u_1(4h,x,t)$$
$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}dk\sum_{j=1}^{\infty}\{R(k,s)G(s)e^{i(kx-ist)}\}_{s=\pm i\omega_j(k)}. \tag{13}$$

Here $R(k,s)G(s)$ is the residue of the integrand, where $R(k,s)$ is given by

$$R(k,s)=\frac{N(k,s)}{d(\det\mathbf{M})/ds} \tag{14}$$

and it may be shown to be an odd function of s.

For the step function input $H(t)$, the Laplace transform $G(s)=1/s$ and the transform equation (13) may be written as

$$u_1(4h,x,t)=\frac{i}{\pi}\sum_{j=1}^{\infty}\int_{-\infty}^{\infty}R_j(k)\frac{[1-\cos\omega_j(k)t]}{\omega_j(k)}e^{ikx}\,dk, \tag{15}$$

where

TABLE I. Material constants of fiber composite material.

ρ (Kgm^{-3})	c_{11} (GPa)	c_{12} (GPa)	c_{23} (GPa)	c_{33} (GPa)	c_{55} (GPa)
1578	13.92	6.92	6.44	160.73	7.07

$$R_j(k)=[R(k,s)]_{s=i\omega_j(k)}=-[R(k,s)]_{s=-i\omega_j(k)}. \quad (16)$$

It may further be shown that $R_j(k)$ is an even function of k so that Eq. (15) reduces to

$$u_1(4h,x,t) = \frac{2i}{\pi}\sum_{j=1}^{\infty}\int_0^{\infty} R_j(k)\frac{[1-\cos\omega_j(k)t]}{\omega_j(k)}\cos kx\,dk. \quad (17)$$

III. RESULTS AND DISCUSSION

The material constants employed in order to obtain numerical results are given in Table I. The results are presented as graphs showing the time history of the upper surface normal displacement due to a sharp step crack opening line dislocation given by $u_1(x_1,x,t)=\delta(x_1-rh)\delta(x)H(t)$ acting in turn at each of the interfaces $x_1=rh$ $(r=0,\pm1,\pm2,\pm3)$. The resulting upper surface normal displacement is given by Eq. (17) which involves an infinite sum of infinite integrals and in order to perform the numerical evaluation the range of integration is restricted to a finite interval $(0,\hat{k})$ and the summation is restricted to a finite number of branches of the dispersion equation.

In performing the evaluations, it is convenient to nondimensionalize the distance x and the time t by introducing $X=x/h$ and $T=ct/h$, where c is a typical wave speed. This results in nondimensional transform parameters $K=kh$ and $\Omega=\omega h/c$ and the numerical evaluations have been carried out for $\hat{K}=10$, using a total of 20 branches of the dispersion curves, 10 for each of the symmetric and the antisymmetric modes. In earlier work,[13] it was found necessary to employ a Hamming window to smooth the cutoff at the upper limit of the wave number in order to eliminate the Gibbs oscillation which gave rise to a noncausal precursor. Here the numerical integration has been carried out without any smoothing of the cutoffs, but the results show no trace of contamination by Gibbs oscillation. This is due to the fact that the residues $R_j(k)$ associated with the different modes have each become negligibly small at the cutoff wave number $\hat{k}$. The restriction of the summation to a total of 20 branches is purely empirical, being based on earlier results which appeared to show no difference in response on increasing the range of the summation from 10 to 20. The results which are presented here in fact contradict this earlier experience since they show that in some circumstances the higher harmonics can make substantial contributions to the overall response.

In these calculations $c=2.97$ mm/μs and taking $h=0.125$ mm gives a plate thickness of 1 mm and the numerical integration is thus taken over the range of wave numbers from 0 to $\hat{k}=8\times10^4$ m^{-1}, corresponding to wavelengths varying from ∞ to approximately 75 μm. For these values of c and h, the nondimensional unit of time $T=1$ corresponds to $t=0.042$ μs and the angular frequency ω ranges from zero to a maximum of about 30 MHz. These results are accordingly consistent with the restrictions arising from the assumption that the fiber composite may be modeled as a homogeneous continuum.

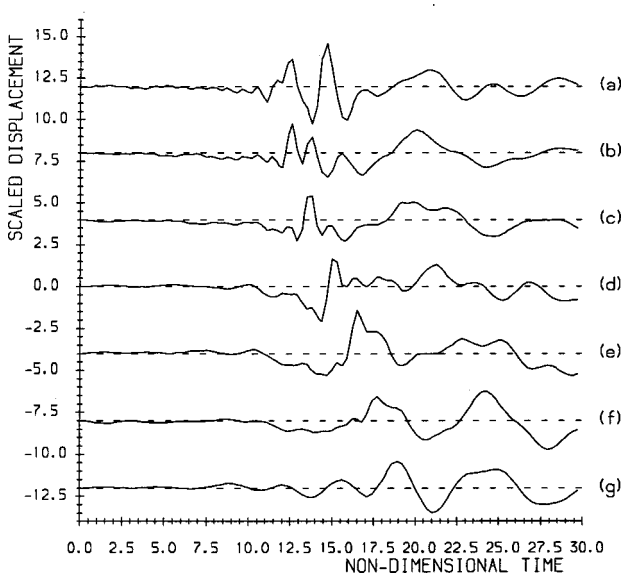


FIG. 1. Upper surface normal displacement at $x=8h$ due to a crack opening dislocation at (a) $x_1=3h$, (b) $x_1=2h$, (c) $x_1=h$, (d) $x_1=0$, (e) $x_1=-h$, (f) $x_1=-2h$, and (g) $x_1=-3h$. Propagation at $\gamma=0$ degrees.

Figures 1–6 each consist of seven curves, each curve showing the time history of the scaled upper surface normal displacement corresponding to a line source acting at one of the seven interfaces between the plies. The zeros of six of the seven curves have been displaced upwards or downwards for clarity of presentation. The graphs are obtained by combining contributions to the expression (17) from each of the first ten modes of both the symmetric and the antisymmetric solutions. In the case of the crack opening dislocation source the symmetric solutions are derived for any location by plac-

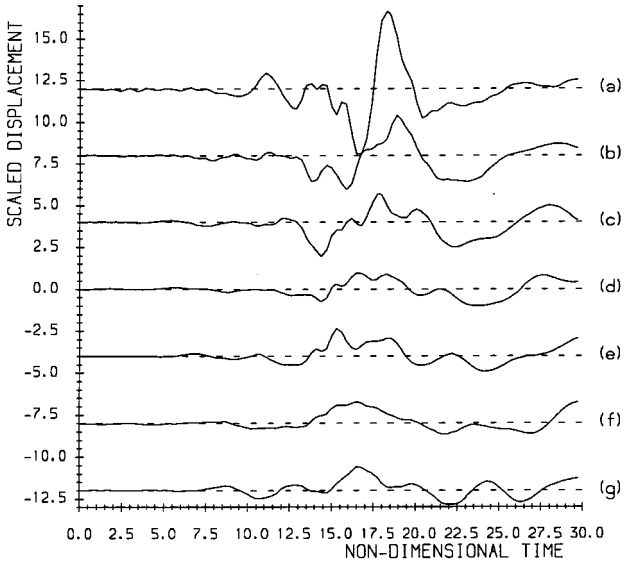


FIG. 2. Upper surface normal displacement at $x=8h$ due to a crack opening dislocation at (a) $x_1=3h$, (b) $x_1=2h$, (c) $x_1=h$, (d) $x_1=0$, (e) $x_1=-h$, (f) $x_1=-2h$, and (g) $x_1=-3h$. Propagation at $\gamma=90$ degrees.

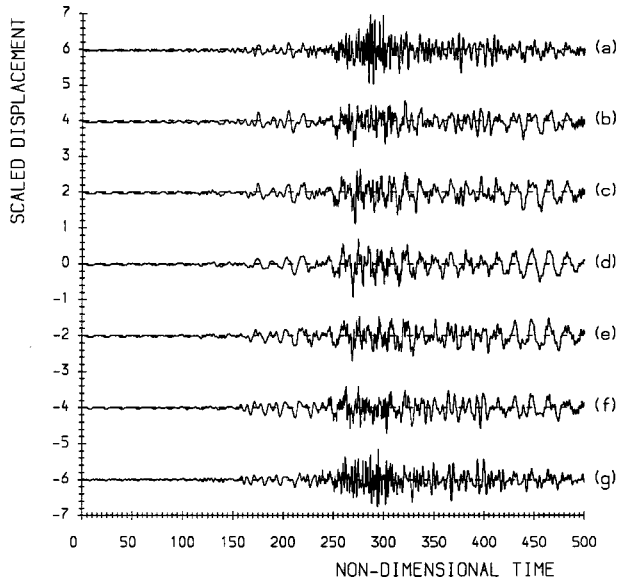


FIG. 3. Upper surface normal displacement at $x=160h$ due to a crack opening dislocation at (a) $x_1=3h$, (b) $x_1=2h$, (c) $x_1=h$, (d) $x_1=0$, (e) $x_1=-h$, (f) $x_1=-2h$, and (g) $x_1=-3h$. Propagation at $\gamma=0$ degrees.

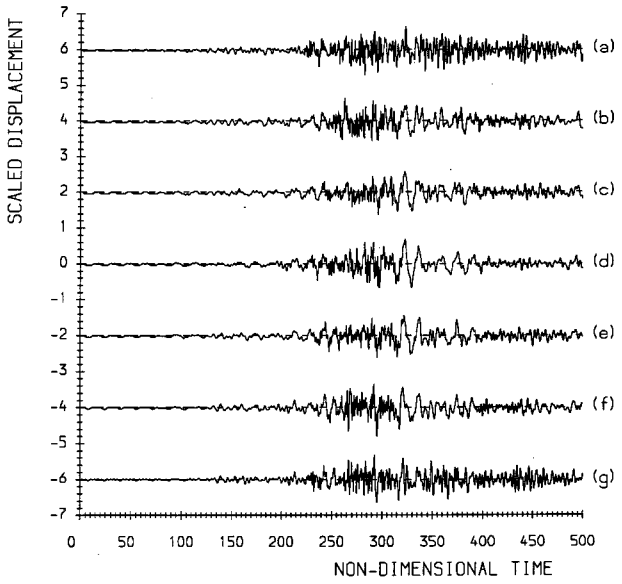


FIG. 5. Upper surface normal displacement at $x=160h$ due to a crack opening dislocation at (a) $x_1=3h$, (b) $x_1=2h$, (c) $x_1=h$, (d) $x_1=0$, (e) $x_1=-h$, (f) $x_1=-2h$, and (g) $x_1=-3h$. Propagation at $\gamma=30$ degrees.

ing a source of one-half the original strength at that location and an identical source at the mirror image location relative to the mid-plane. For the antisymmetric solution a source of one-half the original strength is placed at the source location and an equal and opposite source is placed at the mirror image. In this way it is only necessary to solve for sources at the top three interfaces and at the mid-plane, in order to obtain the seven curves shown in Figs. 1–6. Thus, the top three curves are obtained by adding together the symmetric and antisymmetric solutions, the bottom three by subtracting the antisymmetric solution from the symmetric solution, while the middle curve is derived from the symmetric modes only.

In Fig. 1, the line sources act in the plane $x_3=0$ giving rise to straight crested waves traveling in the direction Ox_3 and the receiver location is at a distance of $8h$ (one plate thickness) from the plane of the sources along the propagation direction, which is the fiber direction in the surface layer. The curves in Fig. 2 represent the response to sources lying in the plane $x_2=0$ and a surface receiver located at a distance of one plate thickness along the propagation direction Ox_2, at right angles to the surface fiber direction. Figures 3 and 4 show the upper surface response to the same sources as Figs. 1 and 2, respectively, when the receiver is located at a distance of $160h$ (20 plate thicknesses) from the plane of the sources. In discussing these figures, it is perti-

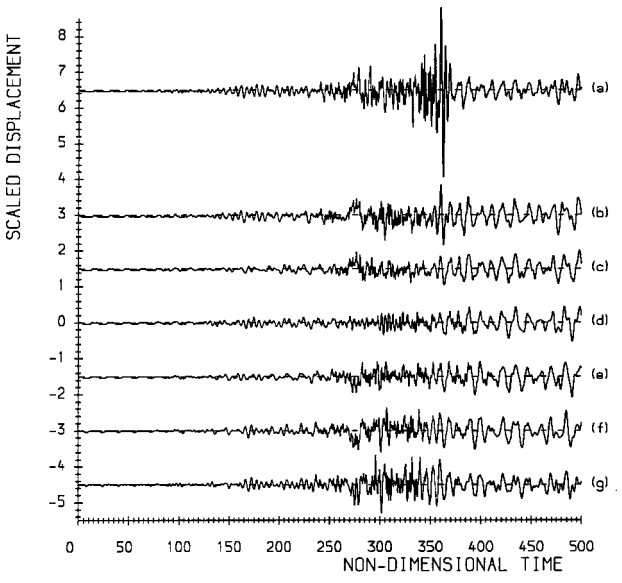


FIG. 4. Upper surface normal displacement at $x=160h$ due to a crack opening dislocation at (a) $x_1=3h$, (b) $x_1=2h$, (c) $x_1=h$, (d) $x_1=0$, (e) $x_1=-h$, (f) $x_1=-2h$, and (g) $x_1=-3h$. Propagation at $\gamma=90$ degrees.

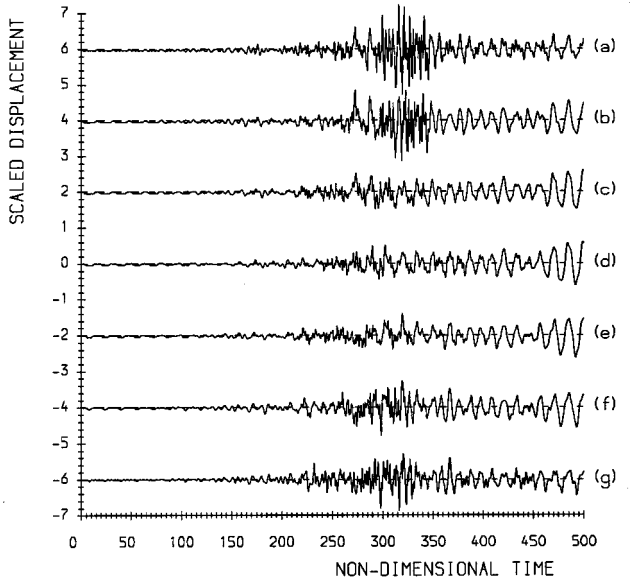


FIG. 6. Upper surface normal displacement at $x=160h$ due to a crack opening dislocation at (a) $x_1=3h$, (b) $x_1=2h$, (c) $x_1=h$, (d) $x_1=0$, (e) $x_1=-h$, (f) $x_1=-2h$, and (g) $x_1=-3h$. Propagation at $\gamma=60$ degrees.

TABLE II. Body and surface wave speeds in fiber composite material.

γ (degrees)	v_1 (mm/μs)	v_2 (mm/μs)	v_3 (mm/μs)	v_R (mm/μs)
0	10.092	2.117	2.117	2.092
15	9.764	2.176	2.084	2.043
30	8.813	2.321	1.981	1.921
45	7.317	2.503	1.832	1.752
60	5.427	2.665	1.670	1.573
75	3.500	2.701	1.540	1.436
90	2.970	2.117	1.489	1.388

nent to note that in an anisotropic material there are three body wave speeds associated with each propagation direction and these are given in Table II for waves propagating in the plane of a ply at different angles γ to the fiber direction. The table also gives the Rayleigh wave speed v_R in each of these directions for surface waves propagating on a half-space of the material.

In the case of Figs. 1 and 2, the receiver is located at a distance of $8.06h$ from the source at the top interface and at distances of $8.25h$, $8.54h$, $8.94h$, $9.43h$, $10.00h$, and $10.63h$ from the sources at each of the other interfaces in order of increasing depth. For the response shown in Figs. 3 and 4, the distance from the source to the receiver varies from a minimum of $160.0h$ to a maximum of $160.15h$. Not surprisingly, the seven curves shown in Fig. 1 differ significantly from each other, reflecting the variation in the distance of travel from source to receiver, and the same observation holds for Fig. 2. In both these figures, the distance of travel is of the order of one laminate thickness and the resulting signals show individual rays arriving either directly from the source or after reflection from one or more interfaces. Nevertheless, the corresponding curves in Figs. 1 and 2 differ considerably from each other. Thus, in Fig. 1 the maximum amplitude decreases gradually and the first arrival time increases gradually with increasing depth of the source, whereas in Fig. 2 there is a rapid drop in the maximum amplitude as the source depth increases and the signals transmitted from the sources located at and below the mid-plane are substantially less than those received from the sources located above the mid-plane. Moreover, the maximum amplitude of each of the top two curves in Fig. 1 is considerably less than that of the corresponding curves in Fig. 2. Both these observations can be accounted for in terms of the material response associated with the directions of the wave propagation in the two cases. In Fig. 1, where the propagation is along the fiber direction in the outer layers, the bending stiffness in the top and bottom layer is high and that in the two middle layers is low, whereas in Fig. 2 it is the bending stiffness of the two middle layers which is high and that of the outer layers which is low. Thus, in Fig. 1 the high stiffness of the outer layer leads to a smaller response from the upper interface source than is the case in Fig. 2 where the outer layer is of much lower stiffness. On the other hand, the low stiffness of the two middle layers in Fig. 1 allows more disturbance to reach the upper surface from the sources in the bottom half of the plate than is the case in Fig. 2, where the middle layers have a much higher stiffness.

Figures 3 and 4 show the response at a distance of 20 plate thicknesses from the planes of action of the sources for propagation along and at right angles, respectively, to the fiber direction in the outer layer. While the responses shown in Figs. 1 and 2 arise mainly from the individual rays, at a distance of 20 plate thicknesses from the source it is to be expected that the disturbances arise from the propagating modes of the Rayleigh–Lamb waves. Here, the variation in distance between the receiver and each of the seven sources is negligible and in Fig. 3 there is a strong similarity between the signals received from sources at corresponding locations on either side of the mid-plane. Thus, graphs (a) and (g) show strong similarities to each other as do graphs (b) and (f) and graphs (c) and (e). This implies that the motion arises predominantly from the symmetric Rayleigh–Lamb modes and this is further supported by the fact that the amplitudes in graph 4, which arises from the mid-plane source, is of the same order of magnitude as in the other graphs. In the case of Fig. 4, the most striking feature is the large amplitude signal to be seen in the top graph at a nondimensional time $T \approx 360$ which corresponds to a disturbance moving with speed 1.32 mm/μs, close to the Rayleigh wave speed. There is evidence of a similar but smaller arrival in the second graph. Aside from these the corresponding graphs on either side of the mid-plane, namely (a) and (g); (b) and (f); and (c) and (e), are virtually mirror images of each other, indicating that in this case the motion is predominantly governed by the antisymmetric Rayleigh–Lamb modes, a conclusion which is further supported by the small amplitude of the disturbance associated with the source at the mid-plane [graph (d)].

In Figs. 5 and 6 the receiver is again at a distance of 20 plate thicknesses from the plane of action of the source. This plane is now aligned at an angle of 60 degrees to the reference direction in Fig. 5 and at an angle of 30 degrees to the reference direction in Fig. 6, giving rise to wave fronts propagating at angles $\gamma = 30$ degrees and $\gamma = 60$ degrees, respectively. The seven curves in each figure no longer show either the symmetry or antisymmetry of Figs. 3 and 4 and each curve is distinct from each of the others. The noticeable feature here is associated with the top two curves in Fig. 6 which both show a high frequency pulse arriving between $T \approx 320$ and $T \approx 360$.

The graphs shown in Figs. 1–6 are all obtained by combining the first ten symmetric and the first ten antisymmetric Rayleigh–Lamb modes of plane sinusoidal wave propagation in the laminate along the appropriate direction. In a plate of isotropic elastic material the short wavelength (high frequency) limiting phase velocity of the fundamental modes of both symmetric and antisymmetric motions tends to the speed of the Rayleigh surface wave, while the limiting phase velocity of all higher modes tend to the shear wave speed in the material. In the case of the quasi-isotropic laminate, the short wavelength limiting behavior of all the modes is angle dependent. For the fundamental modes the limiting speed in any direction is the lower of the Rayleigh wave speed in the outer layer or the slowest body wave speed in each of the inner layers, while for all higher modes the limiting phase velocity is the lowest of the slowest body wave speed in each layer. Figure 7 shows the dispersion curves (scaled phase speed versus dimensionless wave number) for the first three

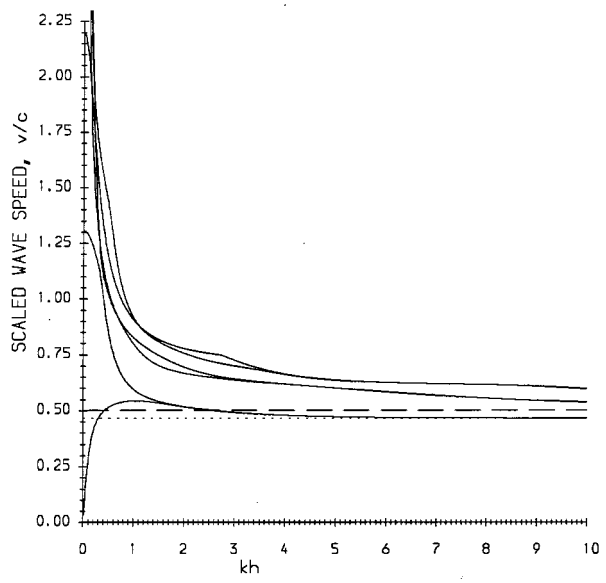


FIG. 7. First six dispersion curves for propagation at $\gamma=90$ degrees. Dashed line is slowest body wave speed in outer layer. Dotted line is surface wave speed in outer layer.

symmetric and the first three antisymmetric modes for propagation at the angle $\gamma=90$ degrees while Fig. 8 shows the corresponding curves for $\gamma=0$ degrees. The data given in Table II show that for propagation at $\gamma=90$ degrees, the Rayleigh wave speed of 1.388 mm/μs and the slowest body wave speed in the outer layer (1.489 mm/μs) are both less than any of the body wave speeds in the internal layers and from the dispersion curves in Fig. 7 it is evident that the two lowest curves tend to the nondimensional surface wave speed (0.467) while the remaining harmonics approach the slowest of the outer layer body wave speeds (nondimensional value 0.501), as the wavelength decreases. This behavior mirrors that of the dispersion curves for an isotropic plate. The existence of the large amplitude disturbance to be

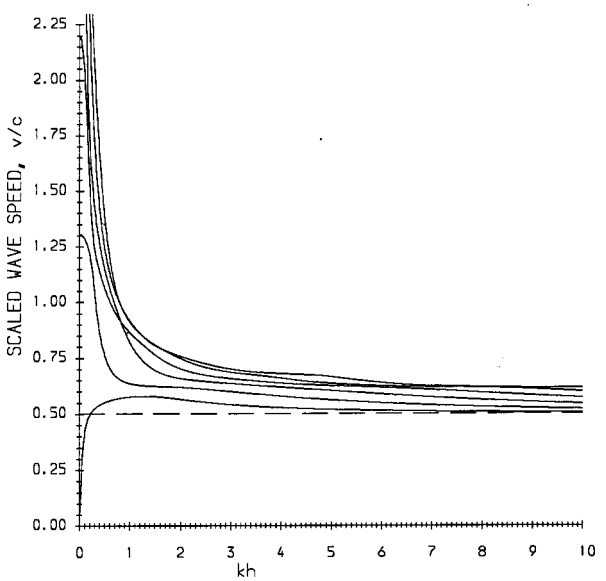


FIG. 8. First six dispersion curves for propagation at $\gamma=0$ degrees. Dashed line is slowest body wave speed in core layers.

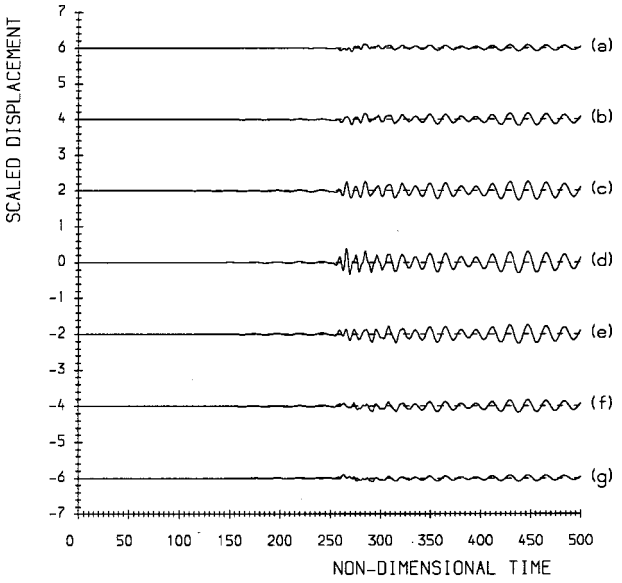


FIG. 9. Contribution from modes 1 to upper surface normal displacement at $x=160h$ due to a crack opening dislocation at (a) $x_1=3h$, (b) $x_1=2h$, (c) $x_1=h$, (d) $x_1=0$, (e) $x_1=-h$, (f) $x_1=-2h$, and (g) $x_1=-3h$. Propagation at $\gamma=0$ degrees.

seen in the top curve of Figs. 2 and 4 is associated with the surface wave arising from the fundamental modes.

In Fig. 8 all six dispersion curves tend towards the same nondimensional wave speed 0.501 as k increases. This reflects the fact that at this orientation the limiting phase velocity of all modes is 1.489 mm/μs, namely the slowest body wave speed in the two inner layers of the laminate. For propagation at $\gamma=60$ degrees, the angle between the fiber direction and the propagation direction in the second layer from the top and bottom is 75 degrees and the slowest body wave speed in this layer (1.540 mm/μs) is lower than both the surface and slowest body wave speeds in the outer material (1.573 and 1.670 mm/μs, respectively). In consequence all of the Rayleigh–Lamb modes tend to this speed in the short wavelength limit. Finally, for propagation at $\gamma=30$ degrees the limiting speed of all modes is again 1.540 mm/μs, which is the slowest body wave speed in the third layer from the top and bottom of the laminate.

It is of interest to examine the contributions from the individual Rayleigh–Lamb modes to the overall responses shown in Figs. 3–6. Figures 9 and 10 show the contributions to the graphs shown in Figs. 3 and 4, respectively, which arise from the fundamental modes of symmetric and antisymmetric motion only. In Fig. 9, it is evident that the fundamental modes make little contribution in the case of the sources at the top and bottom interfaces with the contribution increasing with the depth of the source. Figure 10 shows substantial contributions arising for each source and the top curve clearly shows the large amplitude surface wave motion which is associated with the limiting high frequency phase speed at this orientation. Figures 11 and 12 show the contributions to Figs. 3 and 4, respectively, arising from the combined effect of the seventh mode of both symmetric and antisymmetric motion. From Fig. 11 it may be seen that in the case of the sources at the top and bottom interfaces, these

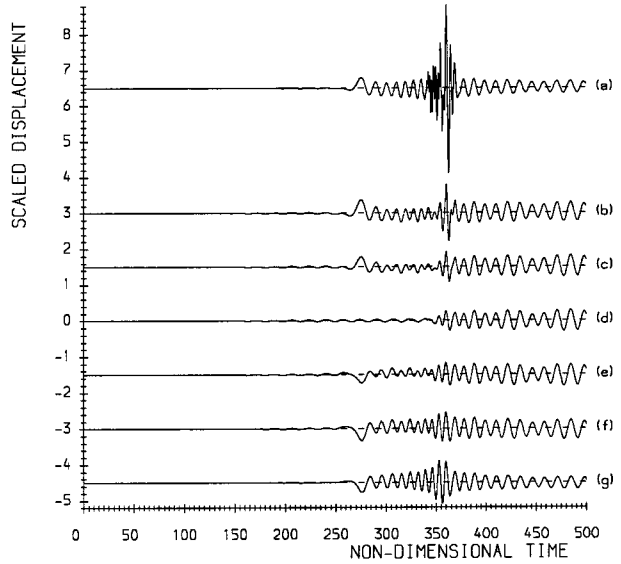


FIG. 10. Contribution from modes 1 to upper surface normal displacement at $x=160h$ due to a crack opening dislocation at (a) $x_1=3h$, (b) $x_1=2h$, (c) $x_1=h$, (d) $x_1=0$, (e) $x_1=-h$, (f) $x_1=-2h$, and (g) $x_1=-3h$. Propagation at $\gamma=90$ degrees.

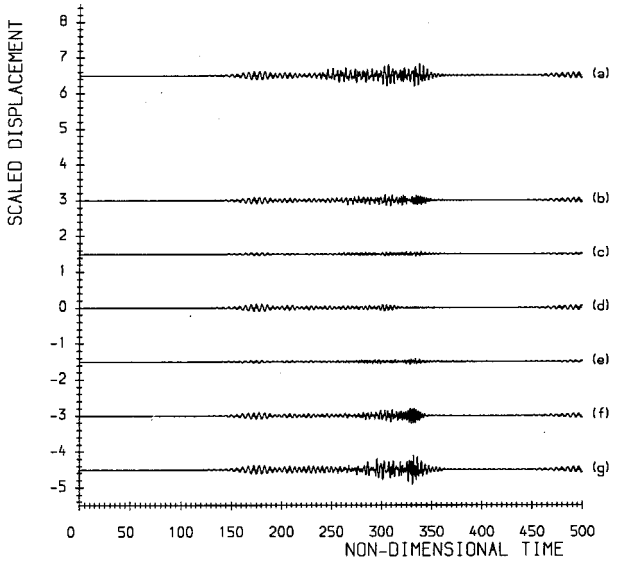


FIG. 12. Contribution from modes 7 to upper surface normal displacement at $x=160h$ due to a crack opening dislocation at (a) $x_1=3h$, (b) $x_1=2h$, (c) $x_1=h$, (d) $x_1=0$, (e) $x_1=-h$, (f) $x_1=-2h$, and (g) $x_1=-3h$. Propagation at $\gamma=90$ degrees.

modes make a significant contribution to the overall response, with a decreasing contribution with increasing depth of source. The seventh modes make much smaller contributions for propagation at $\gamma=90$ degrees as seen in Fig. 12.

IV. CONCLUSIONS

The results presented here show that the transform method has the capability to predict acoustic emission signals in both the near field and the far field. At some orientations, the received signals arising from internal sources located at each interface are sufficiently different from each other to allow the possibility of determining the level at which the delamination occurs. At other orientations, the signals arising from sources located at any pair of interfaces which are symmetrically placed relative to the mid-plane are virtually indistinguishable.

The appearance or nonappearance of surface waves is goverened by the limiting phase velocity of the fundamental Rayleigh–Lamb modes of symmetric and antisymmetric motion. This in turn is related to the orientation of the load line and to the lay-up of the plies. Details of the variation of displacement with position through the laminate are discussed in Ref. 14.

Significant contributions to the surface signal can arise from higher modes of the dispersion equation and approximate solutions based on classical plate theory will fail to include these contributions. It may well be that it is these contributions which will serve to distinguish sources located at different interfaces.

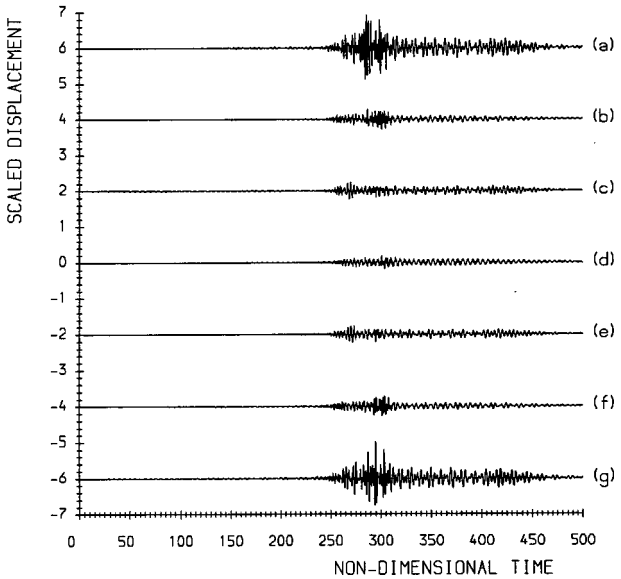


FIG. 11. Contribution from modes 7 to upper surface normal displacement at $x=160h$ due to a crack opening dislocation at (a) $x_1=3h$, (b) $x_1=2h$, (c) $x_1=h$, (d) $x_1=0$, (e) $x_1=-h$, (f) $x_1=-2h$, and (g) $x_1=-3h$. Propagation at $\gamma=0$ degrees.

[1] H. Suzuki, M. Takemoto, and K. Ono, ''A study of fracture dynamics in a model composite by acoustic emission signal processing,'' J. Acoust. Emiss. **11**, 117–128 (1993).

[2] D. Guo, A. K. Mal, and K. Ono, ''Wave theory of acoustic emission in composite laminates,'' J. Acoust. Emiss. **14**, S19–S46 (1996).

[3] S-S. Lih and A. K. Mal, ''On the accuracy of approximate plate theories for wave field calculations in composite laminates,'' Wave Motion **21**, 17–34 (1995).

[4] S-S. Lih and A. K. Mal, ''Response of multilayered composite laminates to dynamic surface loads,'' Composites, Part B **27**, 633–641 (1996).

[5] E. R. Green, ''Acoustic emission sources in a cross-ply laminated plate,'' Composites Eng. **5**, 1453–1469 (1995).

[6] E. R. Green and W. A. Green, ''Predicted far field response of quasi-isotropic laminates to internal impulsive events,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1999), Vol. 18, p. 1225.

[7] D. E. Chimenti, ''Guided waves in plates and their use in materials characterization,'' Appl. Mech. Rev. **50**, 247–284 (1997).

[8] E. R. Green and W. A. Green, ''A computational method for wave propagation from a point load in an anisotropic material,'' Ultrasonics **38**, 262–266 (2000).

[9] E. R. Green and W. A. Green, ''Two dimensional wave propagation from a point load in an anisotropic elastic material,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (AIP, New York, 2000), Vol. 19, p. 137.
[10] A. K. Mal, Y. Bar-Cohen, and S-S. Lih, ''Wave attenuation in fiber-reinforced composites,'' in *M^3D ASTM STP 1169*, edited by V. K. Kinra and A. Wolfenden (1992), p. 245.
[11] A. K. Mal, ''Wave propagation in layered composite laminates under periodic surface loads,'' Wave Motion **10**, 257–266 (1988).
[12] A. K. Mal, ''Guided waves in layered solids with interface zones,'' Int. J. Eng. Sci. **26**, 873–881 (1988).
[13] E. Rhian Baylis and W. A. Green, ''Impact stress waves in fiber-reinforced laminated plates,'' in Proc. 3rd Int. Conf. on Recent Advances in Structural Dynamics, edited by M. Petyt, H. F. Wolfe, and C. Mei, AFWAL-TR-88-3034 (1988), p. 171.
[14] E. R. Green and W. A. Green, ''Channelling of elastic waves in an eight-ply fiber composite plate,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (AIP, New York, 2000), Vol. 19, p. 1143.

The effect of moisture on compressional and shear wave speeds in unconsolidated granular material

F. Douglas Shields and James M. Sabatier
National Center for Physical Acoustics, University of Mississippi, University, Mississippi 38677

Mark Wang
Sensys Instruments Corporation, 1200 Crossman Avenue, Suite 103, Sunnyvale, California 94089

(Received 14 February 2000; accepted for publication 7 August 2000)

The effect of water vapor on the shear and compressional wave speeds in two different kinds of glass beads and in Ottawa sand has been measured. The nominal diameter of the glass beads was 125 μm and of the sand, 500 μm. Measurements were made as the water vapor was introduced slowly into the evacuated material. The vapor pressure isotherm for the beads made of glass with a high concentration of titanium and barium oxides was fit reasonably well by the simple Brunauer–Emmett–Teller (BET) theory. For the Ottawa sand, the BET theory fit the vapor pressure isotherm if the surface area of the grains was assumed to be three times the area calculated, assuming all of the grains were spheres with a diameter of 500 μm. In these two materials, the vapor had little effect on the wave speeds. For beads made of glass containing sodium oxide, however, the wave speeds approximately double with the introduction of water vapor, and the vapor pressure isotherm had the BET shape only if the saturated vapor pressure was assumed to be lowered by 20%. These results have been explained by assuming that a chemical reaction occurred between the lime glass and the water to form a gel. © *2000 Acoustical Society of America.* [S0001-4966(00)02211-6]

PACS numbers: 43.20.Jr [DEC]

I. INTRODUCTION

Water introduced into unconsolidated granular materials has some well-known effects but some of these effects are little understood. For example, moist sand has a completely different consistency from dry sand. (Sand castles can be made with moist sand but not with dry.) This difference would lead one to expect compressional and shear wave velocities to be very different in the moist and dry material. Numerous measurements in the past have shown that this is not the case.[1–4] Furthermore, water-saturated sand is different from either moist or dry sand. The compressional wave velocity in water-saturated sand is generally about 20% greater than in water (i.e., about 1800 m/sec).

However, the completely saturated condition is not easily obtained. Sand must be thoroughly stirred into the water before air can be eliminated and the velocity raised to 1800 m/sec. Even a few tenths of a percent of air in the material causes the compressional velocity to drop to a value close to its value in the dry material (i.e., about 200 m/sec). This small amount of air is sufficient to make the interstitial fluid ''soft'' and the elastic properties of the bulk material are then controlled by the ''frame'' moduli.

In unconsolidated granular material, these frame moduli are determined by the ''stiffness'' of contact points between the grains. (The contact stiffness is defined as the fractional change in separation between centers of touching particles divided into the force causing the change.) This paper presents the results of measurements of the effect of water vapor on the compressional and shear wave velocities in three different granular materials and seeks to explain the results in terms of contact mechanics. The water vapor was introduced after the materials had been evacuated to a few millitorr of pressure. The vapor pressure and the compressional and shear wave velocities were measured as a function of the amount of water introduced. The three materials studied were sodium–silicon (soda-lime) glass beads, titanium–barium glass beads, and ''Ottawa'' sand. Table I gives the chemical composition and some of the physical properties of the samples. Figure 1 shows scanning electron microscopy (SEM) photos of a sample of the glass beads. Figure 2 shows a similar photo of a sample of the sand.

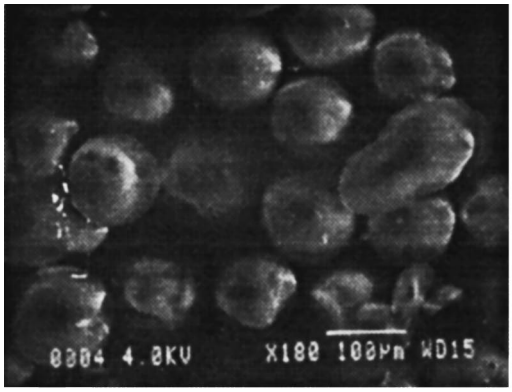
The elastic properties given in this table are approximate. The shear modulus and Poisson's ratio for the Na–Si beads are those given for soda-lime glass in Ref. 5 with the shear modulus calculated from Young's modulus and Poisson's ratio. The values for sand are those given in the AIP Handbook[6] for fused silica and the values given for the Ti–Ba beads are estimates. Properties for the Ti–Ba beads are unavailable because this material is difficult to obtain in the amorphous state except when it is cast in the form of the beads.

II. EXPERIMENT

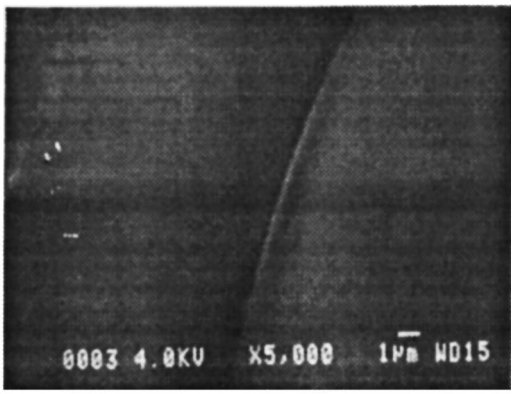
The experimental apparatus is described in detail in Yunlong Wang's dissertation.[7] For the measurements reported here, two different size cylindrical tanks were used. For the sand and the Na–Si beads the diameter was 45.7 cm. For the Ti–Ba beads the diameter was 25.4 cm. The depth at which the compressional velocities were measured in the three materials was approximately 34 in sand, 52 in the Na–Si beads, and 31 cm in the Ti–Ba beads. These depths correspond to overburden pressures of 6.0 for sand, and 8.1 kPa for the two types of beads. These pressures neglect any correction for the part of the weight of the material supported

TABLE I. Properties of samples.

Compound	Percent composition Na–Si glass	Ti–Ba glass	Sand
SiO_2	71–74	11.1	99.8
Na_2O	12–15		
CaO	8–10		
TiO2		43.4	
BaO		41.1	
Other	1–6	4.4	0.2
	Physical property		
Shear modulus (Pa$\times 10^{-10}$)	2.75	2.0	3.1
Poisson's ratio	0.23	0.2	0.17
Specific gravity	2.5	4.2	2.65
Porosity (percent)	36	37	33
Nominal diameter (mm)	0.125	0.125	0.50
Calculated dry *Vp* (m/sec)	285.0	196.0	265.0
Measured dry *Vp* (m/sec)	285.0	179.0	220.0

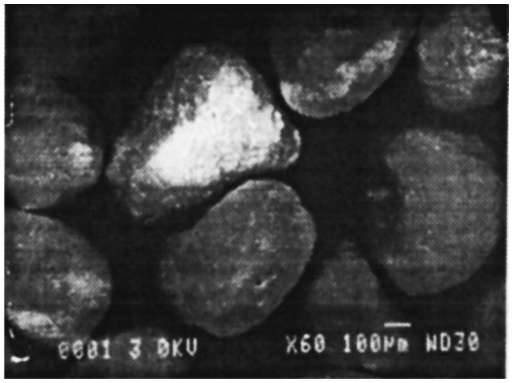


A

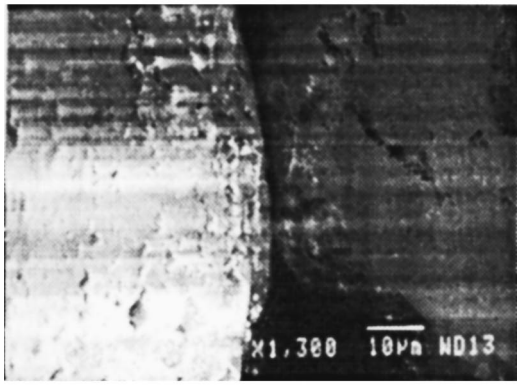


B

FIG. 1. Scanning electron microscope photograph of the glass beads (a) 400 magnification, (b) 1700 magnification.



A



B

FIG. 2. Scanning electron microscope photograph of the Ottawa sand. (a) 60 magnification, (b) 1300 magnification.

by the tank walls (the ''Janssen effect''). However, this correction is small as long as the depth in the material is not much greater than the tank diameter, and this condition is fulfilled here.

A. Transmitters and receivers

The transducers were buried in the material along the axis of the tank 5-cm apart. The compressional wave source was mounted below the receiving transducers and the shear wave source above them. The transducer used to generate compressional waves was made from the sensing element taken from a Teledyne Exploration hydrophone. This element consists of a pair of piezoelectric bimorph discs formed into a hollow capsule about 1.8 cm in diameter. When driven electrically, this capsule radiates compressional waves as a monopole. It was potted in polyurethane so as to be unaffected by the water vapor and the vacuum. The transducer used to generate shear waves was an L410 horizontal geophone 2.2 cm in diameter and 3.3-cm long manufactured by the Mark Products Company. Tests showed it was capable of generating shear waves with frequencies from 700 Hz to 4 kHz without having serious distortion. It had a standard fre-

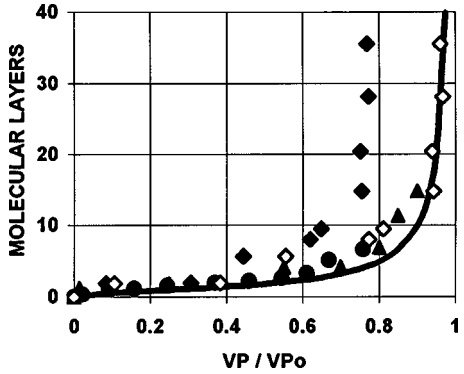


FIG. 3. Surface coverage versus reduced vapor pressure. ♦ Na–Si beads, ● Ti–Ba beads, ▲ sand, ◇ Na–Si beads versus reduced vapor pressure divided by 0.8, — BET theory.

quency of 10 Hz and a coil resistance of 330 ohms. All receiving transducers were ACH-01 accelerometers manufactured by AMP Inc. These accelerometers had dimensions of approximately $1.5\times1.3\times0.4$ cm. They responded to the compressional or shear wave depending upon whether they were oriented parallel or perpendicular to the direction of propagation of the waves. Velocities were determined from measurements of the time of flight for three cycle tone bursts between the receiving accelerometers. The tone burst frequencies were 10 kHz for the compressional waves and 3 kHz for the shear waves. However, the velocities varied little with frequency.

B. Experimental procedure

Before beginning the measurements, the vacuum leak rate of the measuring tank was checked to make sure that a negligible amount of air would leak into it over the several days involved in the measurements. Before the sample was placed in the tank careful measurements were made of the vapor pressure as a function of the amount of water introduced into the empty tank. Then the sample was placed in the tank with careful tamping and evacuated to a few millitorr pressure. When the sample was fully evacuated so as to have a negligible outgassing rate, water vapor was introduced a little at a time. After admitting a small amount of water, the vapor pressure was allowed to stabilize and measurements were made of the compressional and shear velocities.

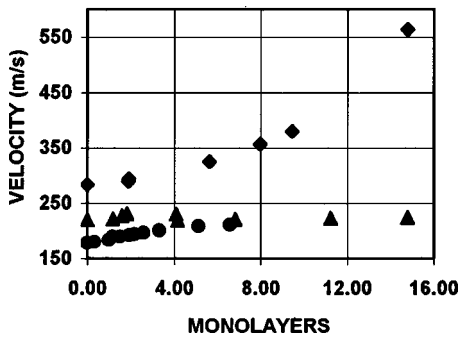


FIG. 4. Compressional velocities versus monolayers of water. ♦ Na–Si beads, ● Ti–Ba beads, ▲ sand.

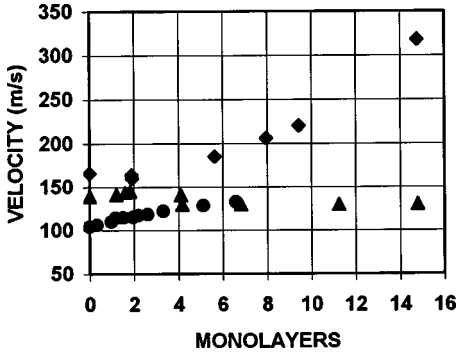


FIG. 5. Shear velocities versus monolayers of water. ♦ Na–Si beads, ● Ti–Ba beads, ▲ sand.

III. RESULTS

The results of the experiment are shown in Figs. 3–7. Figure 3 shows the amount of water adsorbed by the different samples as a function of the reduced vapor pressure. The amount of adsorbed water is given in monolayers of surface coverage. To determine this quantity, the measured amount of water admitted into the sample was first corrected for the amount adsorbed by the tank walls. This correction was determined by measuring the vapor pressure as a function of the mass of water introduced into the empty tank. Then the thickness of the adsorbed layer was calculated by dividing the amount of water taken up by the sample by its surface area. For the beads, this surface area was calculated assuming all of the beads were spheres with a diameter of 125 μm. For the sand, in order to make the resulting curve better fit the theory discussed below, the area was assumed to be three times the area obtained assuming all of the particles were spheres with a diameter of 500 μm. This assumption seemed reasonable since the sand particles were nonspherical, had a range of sizes, and had a much rougher surface than the glass beads. (See the photos in Fig. 2.) The number of monolayers in the coverage was obtained by dividing the thickness of the adsorbed layer by the cube root of the volume of one molecule of water in the liquid state. The reduced vapor pressure

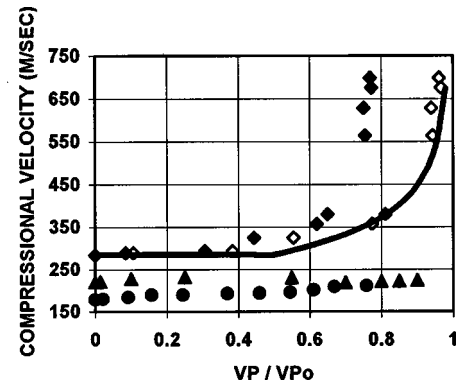


FIG. 6. Compressional velocities versus reduced vapor pressure. ♦ Na–Si beads, ● Ti–Ba beads, ▲ sand, ◇ the measured velocity in the Na–Si beads plotted versus the measured reduced vapor pressure divided by 0.8. The solid curve is a plot of velocity in the sodium beads as explained in the text assuming that the water condensing in the points of contact reacts to form a very high viscosity liquid or solid.

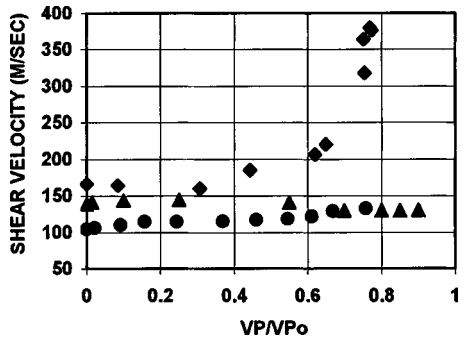


FIG. 7. Shear velocities versus reduced vapor pressure. ♦ Na–Si beads, ● Ti–Ba beads, ▲ sand.

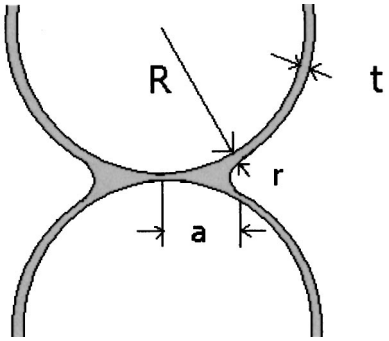


FIG. 8. Diagram of the adsorbed layer and pendular ring of water in the point of contact.

in these figures, p/p_0, is the measured vapor pressure in the sample divided by the vapor pressure of water at the sample temperature.

Figures 4 and 5 show the compressional and shear velocities in the three samples plotted as a function of the adsorbed thickness in monolayers. Figures 6 and 7 show these velocities as a function of reduced vapor pressure. It is obvious from these curves that the effect of the water on the Na–Si glass beads is very different from its effect on the Ti–Ba beads and the sand.

It is also worth noting that $(V_p/V_s)=\sqrt{3}\pm$ about 8%. The value of 3 for $(V_p/V_s)^2$ implies a value of 1/4 for Poisson's ratio for the bulk material. According to the theory by Digby,[8] it also indicates that the tangential stiffness of the contacts between grains is zero. The value of 1/4 for Poisson's ratio is predicted theoretically if one assumes, as Poisson did in his original derivation, that the body is made up of material points that interact with central forces.[9] Doru Velea observed a similar ratio for V_p/V_s in moist Ottawa sand exposed to the atmosphere.[4]

One further interesting experimental result should be noted. With sand and Ti–Ba beads, most of the water introduced into the beads could be relatively quickly pumped out with a corresponding drop in vapor pressure. With the Na–Si beads, the vapor pressure dropped to a few microns when only about 41% of the water was pumped out and the large change in velocity was *not* reversed.

A. Discussion of the results

1. The adsorption isotherms

The adsorption isotherms plotted in Fig. 3 can give some useful information about what is happening to the water that is condensing in the material. Adsorption isotherms are conventionally explained using the BET theory, while Kelvin's equation is usually assumed to give the pressure of a vapor above a curved liquid surface.[10] We can use these two simple theories to give us some idea about how much water is adsorbed on the surface and how much is condensed into the points of contact.

First, consider the volume of water in the pendular rings at the points of contact (Fig. 8). Assuming spherical grains with a radius R, and that the water has a zero angle of contact with the glass, to a good approximation, the volume of water in the contact divided by the volume of the bead is given by

$$\frac{V}{V_0}=\frac{3}{4}\left(\frac{r}{R}\right)^2, \tag{1}$$

where R is the radius of the bead and r is the radius of curvature of the water surface. This curvature is related to the vapor pressure by Kelvin's equation,[10]

$$r=\frac{s\nu_1}{R_0T}\ln\left(\frac{p}{p_0}\right). \tag{2}$$

Here, R_0 is the gas constant, T is absolute temperature, v_l is the specific volume of the liquid, s is the surface tension, p is the vapor pressure above the curved surface, and p_0 is the vapor pressure above a flat surface at the same temperature. The equation obtained by substituting Eq. (2) into Eq. (1) gives the amount of water in the points of contact as a function of the vapor pressure.

The amount of water adsorbed on the surface can be estimated from the BET theory. According to that theory,[10]

$$\frac{n}{n_m}=\frac{c(p/p_0)}{(1-p/p_0)[1+(c-1)(p/p_0)]}, \tag{3}$$

where n is the moles of vapor adsorbed, n_m is the moles to cover a monolayer. Therefore, n/n_m is equal to the monolayer coverage, $c=e^{(Q_l-Q_v)/RT}$, where Q_l is the heat of adsorption on the bare surface, and Q_v is the heat of condensation. For water on glass, c is greater than 10, and n/n_m depends little on c.

Using Eqs. (1) and (2) to calculate the water condensed into the points of contact and Eq. (3) to calculate the amount adsorbed on the surface shows that, for any reasonable number of contacts per bead, a very small amount of the total water adsorbed goes into the points of contact. Therefore, Eq. (3) has been used to calculate the theoretical curve in Fig. 3. The experimental measurements for the Ti–Ba beads and the sand agree reasonably well with the BET curve. However, for the Na–Si beads the vapor pressure appears to be approaching a value that is about 80% of that of water.

A similar lowering of vapor pressure was observed in the late 1960's for water in fine capillaries.[10] This observation was first heralded as a sensational discovery of a new form of water that was named ''polywater'' or ''anomalous water,'' or ''cyclimetric water.'' In addition to having a lower vapor pressure, this new kind of water was observed to have a much higher viscosity than regular water. After extensive research and many articles extending over more than

seven years, the discoverers finally were convinced that they were observing a mixture of colloidal and molecularly dissolved impurities that resulted from a reaction between the water and the glass.[10]

If such a reaction is taking place between water and the Na–Si glass, it might explain why the adsorption isotherm and velocity curves for this material differ from those for the Ti–Ba glass and the sand. To test this hypothesis, the Na–Si measured values in Fig. 3 have been plotted against the reduced vapor pressure assuming that the saturation vapor pressure, VPo, has been reduced by 20%. This is seen to bring the measurements on the Na–Si beads in line with those on the Ottawa sand and the Ti–Ba beads, and the BET theory.[10]

With this explanation of where the water is going, we now look at the velocity curves to see what they can tell us about what the water is doing to the elastic properties of the lattice. Wang[7] showed that, because of its relatively small viscosity, the water would flow in and out of the contact during the acoustic cycle and have little effect on the stiffness and, therefore, the wave velocities. In fact, this is what is observed for the effect of water on the sand and the Ti–Ba beads. However, the large increase in velocity in the Na–Si glass beads implies an increase in viscosity that would lock the water in the points of contact. This increase in viscosity was also a property of the bogus ''anomalous water.'' [10] The difference between the effect of the water on the Ti–Ba beads and sand and the Na–Si glass is explained by assuming a chemical reaction between the water and the sodium in the Na–Si glass that produced a ''hydrogel'' with a high viscosity, and a lower vapor pressure. Such a chemical reaction evidently did not take place between the water and the Ottawa sand or the Ti–Ba beads because of the absence of sodium in these materials. An explanation of a possible chemical reaction is as follows:

According to the manufacturer's analysis, the Na–Si beads contain SiO_2, Na_2O, CaO, MgO, Al_2O_3, and other oxides. The composition, however, is an amorphous fusion of these oxides and is not a crystalline type of free oxide (like silica). A silicate like albite, $NaAlSi_3O_8$, could have been one of the mineral compounds in the beads. $NaAlSi_3O_8$ can also be represented by $Na_2SiO_3 \square Al_2O_3 \square 5SiO_2$, that is the mixed formulation of sodium metasilicate, Na_2SiO_3, and oxides. The Na–O bond is a fairly weak bond compared with the Si–O bond or even the Ca–O bond. When Na_2SiO_3 is dissolved in water to form ''waterglass,'' the dissolution may be considered to proceed via the reaction:[11]
$Na_2SiO_3 + 3H_2O = 2NaOH + H_4SiO_4$.

The first product is what makes the glass beads show alkalinity when fine size beads are exposed to water. The second product, more appropriately written in the form of $Si(OH)_4$, is called monosilicic acid, and is a tetrahedral molecule consisting of a silicon atom bonded to four OH groups. In Fig. 9, $Si(OH)_4$ tetrahedra are shown schematically. The squares are silicon atoms and the open circles are OH radicals. In the solution, two molecules of monosilicic acid may combine to form a larger molecule represented by the structure at the bottom right of the figure. In this case, two OH groups of nearby $Si(OH)_4$ molecules have reacted to form a Si–O–Si bridge (indicated by the black circles), with the release of a water molecule.

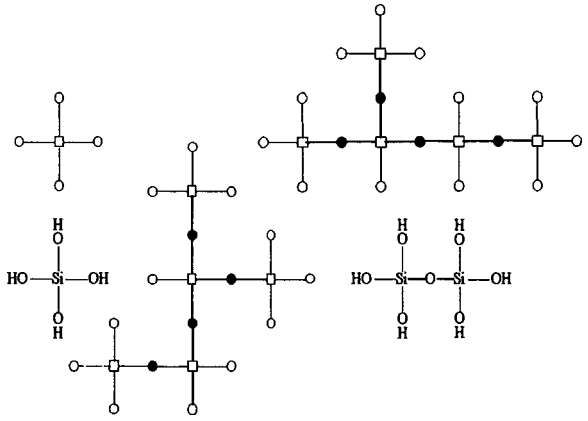


FIG. 9. Polymerization scheme for reaction between the water and the sodium glass. The open squares represent silicon atoms. The open circles represent OH radicals, and the closed circles represent oxygen atoms. The polymerization proceeds by bridging silicon atoms together with an oxygen bond with the liberation of a water molecule.

This polymerization reaction can continue, building up larger molecules as indicated at the top right of Fig. 9. Each initial $Si(OH)_4$ monomer can form as many as four ''reacted bonds,'' that is, oxygen bridges to silicon atoms originating from other molecules. Although in reality the polymerization process is known to occur through the mediation of ionic variants of these species, and not through the simple combining of uncharged silicic acid molecules, the picture in Fig. 9 captures the essential feature of this condensation phenomenon. Because of the stability of the silicon–oxygen bond, this particular condensation process may be viewed as irreversible. Once the Si–O–Si cross links form, they remain.

As the reaction goes on and more cross links form, larger and larger molecules appear. Eventually and abruptly, an infinitely extended[12] molecule appears at a critical stage in the condensation process. The macroscopically extended three-dimensional network molecule is called the gel macromolecule.[11] The process of its formation is called gelation, and its abrupt appearance is called the sol–gel transition. The term sol applies to the solution containing only finite molecules, while the term gel applies to the system containing the extended network.

In the sol phase, only atomic-scale molecules are present, and the material is just a conventional liquid. The viscosity of the material increases as the sol–gel transition is approached, and becomes very large when it is close to transition. Treadwell[12] measured the viscosity in a silicic acid sol during the sol–gel transition. He found the viscosity was already at 1683 Pa·s before the sol was gelled. Other physical properties also change.

From this discussion, it is assumed that a reaction with the sodium in the Na–Si glass produced a product on the surface of the beads that had a lower vapor pressure and a higher viscosity than water; two properties that might be inferred from the measurements discussed above. This assumption is further supported by the fact that for the sodium beads only 41% of water introduced into the tank could be pumped back out.

A theory by Digby[8] can be used to calculate the velocities and the effect of the high viscosity fluid on them. He used Hertz–Mindlin theory to calculate elastic moduli in randomly packed spheres. Winkler[12] uses Digby's equations to express the wave speeds. This analysis shows that, if the contact does not support a shear, $(V_p/V_s)^2=3$, and the compressional wave speed squared should vary as the contact radius. As noted previously, our results confirm the $(V_p/V_s)^2=3$ condition in glass beads and moist sand. We therefore assume that the tangential contact stiffness is zero. In this case, Winkler's equation for the compressional velocity squared is[13]

$$V_p^2=\frac{3Ka\mu}{5\pi R\rho(1-v)}, \tag{4}$$

where K is the number of contacts per bead, a is the contact radius, μ is the shear modulus of the glass, R is the particle radius, ρ is the density of the particle material, and v is Poisson's ratio for the particle material. Further, the equation for the contact radius in terms of the normal force, Y, pushing the spheres together, is

$$a_0=\left[\frac{3(1-v)RY}{8\mu}\right]^{1/3}. \tag{5}$$

According to Digby, the equation for the normal force in terms of the pressure is

$$Y=\frac{4\pi R^2}{(1-\alpha)K}\cdot P, \tag{6}$$

where α is the porosity. P is assumed to be the overburden pressure, and is equal to $\rho gh(1-\alpha)$. Using the physical properties and dry velocity for Na–Si glass beads given in Table I, these equations can be combined to give $K=4$ (which, incidentally, is the minimum contact for stability), and $a_0=0.5\,\mu$m. Recent measurements and numerical simulations give a value of $K=6$ for spheres.[14] However, in arriving at this number, it is assumed that all of the contacts experience a normal force caused by the full overburden pressure. Undoubtedly, this force varies from contact to contact. The total value of $K\cdot a_0$ is probably realistic.

The effect of the water is now included by assuming that it increases the contact radius. As long as $(r+t)\ll a\ll R$,

$$a\sim[2R(r+t)]^{1/2}, \tag{7}$$

with r given by Eq. (2) and t equal to the number of molecular layers obtained by Eq. (3) times the molecular diameter. (See Fig. 8.) It is assumed that as long as ''a'' calculated by Eq. (7) is less than ''a_0,'' the water will have little effect on the velocity. Values of r and t calculated by Eqs. (2) and (3) show that this is the case as long as p/p_0 is less than 0.5. For values of p/p_0 above 0.5, contact radii calculated using Eqs. (2), (3), and (7) are substituted into Eq. (4) to obtain the theoretical curve for V_p plotted in Fig. 6. Again, in order to get the agreement between theory and experiment it is necessary to multiply the vapor pressure of water at the sample temperature by 0.8 as was done to get agreement between the BET theory and the measured isotherms. It should be emphasized that this factor of 0.8 is the only adjustable parameter in obtaining this fit to the data, and that factor is confirmed by the ''anomalous water'' experiment of a decade ago.

2. Comparison of the velocities in the dry materials

If one combines Eqs. (4)–(6) above,

$$V_p=0.82\cdot\left[\frac{K\mu h^{1/2}}{(1-v)\rho}\right]^{1/3}\quad(\text{mks units}). \tag{8}$$

This equation assumes that force at the contact points between the beads is normal to the surface, independent of direction and caused by the overburden pressure. In the calculation of the effect of water on the velocity in the Na–Si beads, this equation was used to determine the number of contacts per bead. Using the same number of contacts per bead $(K=4)$ and the physical properties given in Table I, this equation can be used to calculate V_p in the other materials. These calculated values are given along with the measured dry values in Table I. As mentioned earlier, the shear modulus and Poisson's ratio for the titanium–barium beads are only estimates.

IV. CONCLUSIONS

(1) Elastic waves propagated in the frame of granular material are best understood in terms of contact mechanics (Hertz–Mindlin–Digby–Winkler). The wave velocity squared is proportional to the contacts per particle times the radius of the contact.
(2) Moisture has little effect upon the wave velocities unless the moisture reacts with the surface to produce a high viscosity liquid or solid.
(3) The viscosity of the fluid in the contact is more important than is its elastic moduli in determining its effect on the contact stiffness. Unless constrained, the fluid flow in and out of the contact will relieve any stress due to deforming the liquid. The strain in the frame is produced by the deformation of the grains with negligible strain in the material in the contact.
(4) Water in the contact, even in small amounts, eliminates the tangential stiffness that results in $(V_p/V_s)^2\sim 3$.
(5) Water is condensed in the points of contact by surface adsorption and by surface tension. The vapor pressure in the material is controlled by the adsorption process and not by surface tension.
(6) Water can sometimes react with surface of the grains to form a material with a high viscosity and a vapor pressure lower than that of water. The water ''frozen'' in the contacts increases the contact stiffness by increasing the area of contact.
(7) The wave velocities in dry beads under gravitational loading should vary little with the bead radius.

Further experimental work is planned to test these conclusions.

ACKNOWLEDGMENTS

The U.S. Department of Agriculture supported this work. The authors express sincere appreciation to Dr. S. N. Prasad and Dr. Matt J. M. Römkens for several helpful suggestions.

[1] S. E. Elliot and B. F. Wiley, ''Compressional velocities of partially saturated unconsolidated sands,'' Geophysics **40**, 949–954 (1975).
[2] S. N. Domenico, ''Elastic properties of unconsolidated porous reservoirs,'' Geophysics **42**, 1339–1368 (1977).
[3] T. Bourbie, O. Coussy, and B. Zinszner, *Acoustics of Porous Media* (Gulf, Paris, 1987).
[4] Doru Velea, ''The effect of moisture on the propagation of compressional and shear waves in Ottawa sand,'' Ph.D. dissertation, The University of Mississippi, 1998.
[5] *Engineering Materials*, edited by Z. D. Jastrzebski (Van Nostrand, New York, 1950), p. 153.
[6] E. Gray, *American Institute of Physics Handbook*, 2nd ed. (Wiley, New York, 1959), pp. 3–88.
[7] Y. Wang, ''The effect of water vapor on sound propagation in unconsolidated porous materials,'' Ph.D. dissertation, The University of Mississippi, 1997.
[8] P. J. Digby, ''The effective elastic moduli of porous rocks,'' J. Appl. Mech. **48**, 803–808 (1981).
[9] S. D. Poisson, ''Mémoire sur l'équilibre et le mouvement des corps élastiques,'' Mém. de l'Acad. **8** (1829).
[10] A. W. Adamson, *Physical Chemistry of Surfaces*, 4th ed. (Wiley, New York, 1982), pp. 266, 537, and 253.
[11] R. Zallen, *The Physics of Amorphous Solids* (Wiley, New York, 1983).
[12] W. D. Treadwell, Trans. Faraday Soc. **31**, 297 (1935).
[13] K. W. Winkler, ''Contact stiffness in granular porous materials: Comparison between theory and experiment,'' Geophys. Res. Lett. **10**, 1073–1076 (1983).
[14] H. A. Makse, N. Gland, D. L. Johnson, and L. M. Schwartz, ''Why effective medium theory fails in granular materials,'' Phys. Rev. Lett. **83**, 5070 (1999).

On ultrasonic guided waves in a thin anisotropic layer lying between two isotropic layers

A. Jonas Niklasson,[a] Subhendu K. Datta, and Martin L. Dunn
University of Colorado, Department of Mechanical Engineering, Campus Box 427, Boulder, Colorado 80309-0427

(Received 22 November 1999; accepted for publication 7 August 2000)

In this paper, dispersion of guided waves in a three-layered sandwich plate is considered. The focus here is on a configuration consisting of a thin anisotropic layer sandwiched between two identical isotropic layers. This configuration could model, for example, a superconducting tape where the middle layer is a brittle superconductor and the surrounding layers are isotropic and ductile. An approximate dispersion relation correct to $O(h)$ governing the guided waves is obtained by expanding the field inside the thin middle layer in powers of the small thickness h of the layer. Numerical examples are given for two specific systems with superconducting middle layers. Some characteristic features, particularly at low frequencies, are investigated. Comparison between the exact and approximate dispersion relations are made to show that the approximation works well in the frequency interval of interest. The characteristic features may be useful for ultrasonic measurements of the anisotropic elastic constants of the thin layer. © *2000 Acoustical Society of America.* [S0001-4966(00)02111-1]

PACS numbers: 43.20.Mv, 43.35.Cg [DEC]

I. INTRODUCTION

In this paper, our attention is focused on guided waves in a plate with a *thin* superconducting layer sandwiched between two identical isotropic layers. The motivation for this study is to develop a fundamental understanding of ultrasonic guided waves in tapes that are fabricated for commercial high-current applications. Fabrication processes are being developed for different classes of high-T_c materials systems. The tapes are: (1) in the form of a superconducting oxide layer (such as BSCCO) sandwiched in a high-conductivity silver, (2) dip-coated silver tapes in which the oxide layers are outside the core conducting layer, (3) silver-sheathed dip-coated tapes, and (4) deposition of biaxially aligned oxide (YBCO) coatings on nickel tapes. Mechanical processes such as rolling are used to get the crystallographic texture most favorable to high current capacity of the tapes. Such mechanical processes coupled with thermal cycling cause microcracking of the brittle oxide layer(s), which limits the current carrying capacity. The degree of current carrying capacity reduction is a strong function of the crystallographic texture of the oxide layer and the nature of microcracking, which is also a strong function of the texture. These effects influence the mechanical responses such as ultrasonic velocities and attenuation of guided waves propagating along the tapes. Exploitation of the connection between the electrical and mechanical responses may prove to be an efficient means of nondestructive material property characterization during and after the processing of the tapes.

There is now a large body of literature on ultrasonics in superconducting bulk materials. It is known that for superconductors, the elastic constants can be linked explicitly to the superconducting transition temperature T_c through the Debye temperature Θ_D and the electron-phonon coupling parameter λ.[1,2] A review of various ultrasonic measurements of elastic properties can be found in Ref. 3. Investigation of *in situ* mechanical behavior and properties of thin superconducting layers and tapes has been very limited. Since the properties are highly dependent upon external and internal stress fields, interface properties and porosity, to name a few, it is necessary to understand the basic problem of guided wave propagation in an anisotropic three-layered tape.

Ultrasonic guided wave propagation in thin coating layers on thick elastic substrates or in thin layers sandwiched between thick elastic layers has been investigated by many researchers for applications in, for example, seismology, electronic devices, and thermal barrier coatings. A full review of the literature is beyond the scope of this work. In most of these studies, the substrates or surrounding materials are assumed to be much thicker than the wavelengths of interest so that they can be effectively modeled as semi-infinite media. References to early studies can be found in the review article by Farnell and Adler.[4]

Of particular relevance to the present work are studies of guided waves in anisotropic layered or laminated plates. Considerable progress has been made in recent years on the ultrasonic characterization of anisotropic elastic properties of composite plates. Mathematical modeling and numerical simulation have played a crucial role in the success of experimental methods. This is especially important since the anisotropy has been found to alter the characteristics of guided elastic waves significantly (Auld[5] and Nayfeh[6]). Li and Thompson[7] and Solie and Auld[8] have presented detailed studies of guided waves in homogeneous anisotropic plates. Karunasena *et al.*[9] consider guided waves in multilayered cross-ply composite plates. For an extensive review of the use of guided waves in plates for material characterization,

[a]Current address: Department of Mechanics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden.

we refer to the review articles by Chimenti[10] and Datta,[11] the recent book by Rose,[12] and the references therein.

As mentioned above, the present study is concerned with three-layered superconducting tapes in which the *thin* oxide layer is sandwiched between two metallic layers. The case when the oxide layers are outside the metallic core was studied in an earlier paper.[13] The exact dispersion relation for such multilayered tapes was derived in the latter study and will not be repeated here. Instead, a brief treatment will be given below for the approximate (anisotropic) interface conditions (ICs) when the oxide layer is *thin*. In this paper, the approach taken by Bövik[14] is extended to anisotropic layers. Briefly, the field quantities in the thin layer are expanded in powers of the thickness of the layer in order to derive the interface conditions. This approximation results in a significant simplification of the dispersion equation. Comparison of the approximate and exact dispersion results for two particular cases is shown to be quite good in the frequency range of interest.

II. APPROXIMATE DISPERSION RELATION

A. Interface conditions

Consider a thin anisotropic layer of thickness $2h$ sandwiched between two elastic bodies (see Fig. 1). Let the displacement and stress be denoted by u_j^+ and σ_{jm}^+, respectively, in the lower body ($x_3>h$), by u_j^- and σ_{jm}^-, respectively, in the upper body ($x_3<-h$) and by U_j and Σ_{jm}, respectively, in the layer ($|x_3|<h$). In the layer we assume that the usual elastodynamic equations of motion hold

$$\partial_m \Sigma_{mj} = \rho_L \partial_t^2 U_j, \quad |x_3|<h, \tag{1}$$

where ρ_L is the density of the layer. For simplicity, we assume that the x_1x_2 plane is a plane of elastic symmetry of the layer (i.e., monoclinic material symmetry). Hooke's law is then

$$\begin{pmatrix} \Sigma_{11} \\ \Sigma_{22} \\ \Sigma_{33} \\ \Sigma_{23} \\ \Sigma_{13} \\ \Sigma_{12} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & C_{16} \\ C_{12} & C_{22} & C_{23} & 0 & 0 & C_{26} \\ C_{13} & C_{23} & C_{33} & 0 & 0 & C_{36} \\ 0 & 0 & 0 & C_{44} & C_{45} & 0 \\ 0 & 0 & 0 & C_{45} & C_{55} & 0 \\ C_{16} & C_{26} & C_{36} & 0 & 0 & C_{66} \end{pmatrix} \times \begin{pmatrix} \partial_1 U_1 \\ \partial_2 U_2 \\ \partial_3 U_3 \\ \partial_3 U_2 + \partial_2 U_3 \\ \partial_3 U_1 + \partial_1 U_3 \\ \partial_2 U_1 + \partial_1 U_2 \end{pmatrix}. \tag{2}$$

Since the bonds between the layer and the surrounding bodies are assumed to be perfect, the following conditions must hold:

$$u_j^{\pm} = U_j, \quad \sigma_{j3}^{\pm} = \Sigma_{j3}, \quad x_3 = \pm h. \tag{3}$$

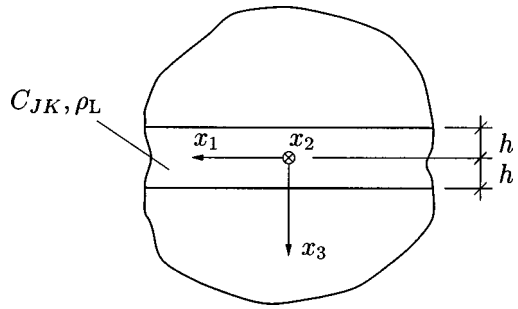


FIG. 1. A thin anisotropic interface layer.

Introduce the notation $(f)_a$ for $f(x_1,x_2,a,t)$ and expand the displacement U_j and traction Σ_{j3} in Taylor series as follows:

$$(U_j)_0 = (U_j)_{\pm h} \mp h(\partial_3 U_j)_{\pm h} + O(h^2), \tag{4a}$$

$$(\Sigma_{j3})_0 = (\Sigma_{j3})_{\pm h} \mp h(\partial_3 \Sigma_{j3})_{\pm h} + O(h^2). \tag{4b}$$

From Hooke's law for the layer, Eq. (2), we easily find the following relations:

$$\partial_3 U_1 = (C_{44}\Sigma_{13} - C_{45}\Sigma_{23})/\delta - \partial_1 U_3, \tag{5a}$$

$$\partial_3 U_2 = (C_{55}\Sigma_{23} - C_{45}\Sigma_{13})/\delta - \partial_2 U_3, \tag{5b}$$

$$\partial_3 U_3 = (\Sigma_{33} - C_{13}\partial_1 U_1 - C_{23}\partial_2 U_2 - C_{36}\partial_2 U_1 - C_{36}\partial_1 U_2)/C_{33}, \tag{5c}$$

where $\delta = C_{44}C_{55} - C_{45}^2$. If Eq. (5) is inserted into Eq. (4a) and the interface conditions (3) are used, we obtain the following relation:

$$(\boldsymbol{u}^+)_h + h(A_u \boldsymbol{u}^+ + A_\sigma \boldsymbol{\tau}^+)_h = (\boldsymbol{u}^-)_{-h} - h(A_u \boldsymbol{u}^- + A_\sigma \boldsymbol{\tau}^-)_{-h} + O(h^2), \tag{6}$$

where $\boldsymbol{u}^{\pm} = (u_1^{\pm}, u_2^{\pm}, u_3^{\pm})^T$, $\boldsymbol{\tau}^{\pm} = (\sigma_{13}^{\pm}, \sigma_{23}^{\pm}, \sigma_{33}^{\pm})^T$, and the nonzero elements of the matrices A_u and A_σ are

$$\begin{aligned} &(A_u)_{13} = \partial_1, \quad (A_u)_{23} = \partial_2, \\ &(A_u)_{31} = (C_{13}\partial_1 + C_{36}\partial_2)/C_{33}, \\ &(A_u)_{32} = (C_{36}\partial_1 + C_{23}\partial_2)/C_{33}, \\ &(A_\sigma)_{11} = -C_{44}/\delta, \quad (A_\sigma)_{22} = -C_{55}/\delta, \\ &(A_\sigma)_{33} = -1/C_{33}, \quad (A_\sigma)_{12} = (A_\sigma)_{21} = C_{45}/\delta. \end{aligned} \tag{7}$$

From the equations of motion, Eq. (1), and Hooke's law, Eq. (2), we get

$$\partial_3 \Sigma_{13} = \rho_L \partial_t^2 U_1 - C_{11}\partial_1^2 U_1 - 2C_{16}\partial_1\partial_2 U_1 - C_{66}\partial_2^2 U_1 - C_{16}\partial_1^2 U_2 - (C_{12}+C_{66})\partial_1\partial_2 U_2 - C_{26}\partial_2^2 U_2 - C_{13}\partial_1\partial_3 U_3 - C_{36}\partial_2\partial_3 U_3, \tag{8a}$$

$$\partial_3 \Sigma_{23} = \rho_L \partial_t^2 U_2 - C_{16}\partial_1^2 U_1 - (C_{12}+C_{66})\partial_1\partial_2 U_1 - C_{26}\partial_2^2 U_1 - C_{66}\partial_1^2 U_2 - 2C_{26}\partial_1\partial_2 U_2 - C_{22}\partial_2^2 U_2 - C_{36}\partial_1\partial_3 U_3 - C_{23}\partial_2\partial_3 U_3, \tag{8b}$$

$$\partial_3 \Sigma_{33} = \rho_L \partial_t^2 U_3 - \partial_1 \Sigma_{13} - \partial_2 \Sigma_{23}. \tag{8c}$$

We now use Eq. (5c) to eliminate the $\partial_3 U_3$ terms from Eq. (8) and insert the result into Eq. (4b) together with the interface conditions (3). We then obtain the condition

$$(\boldsymbol{\tau}^+)_h + h(B_u \boldsymbol{u}^+ + B_\sigma \boldsymbol{\tau}^+)_h = (\boldsymbol{\tau}^-)_{-h} - h(B_u \boldsymbol{u}^- + B_\sigma \boldsymbol{\tau}^-)_{-h} + O(h^2), \tag{9}$$

where $B_\sigma = A_u^T$ and the nonzero elements of the matrix B_u are

$$\begin{aligned}
(B_u)_{11} &= -\rho_L \partial_t^2 + \left(C_{11} - \frac{C_{13}^2}{C_{33}} \right) \partial_1^2 \\
&\quad + 2\left(C_{16} - \frac{C_{13}C_{36}}{C_{33}} \right) \partial_1 \partial_2 + \left(C_{66} - \frac{C_{36}^2}{C_{33}} \right) \partial_2^2, \\
(B_u)_{12} &= (B_u)_{21} \\
&= \left(C_{16} - \frac{C_{13}C_{36}}{C_{33}} \right) \partial_1^2 + \left(C_{12} + C_{66} \right. \\
&\quad \left. - \frac{C_{13}C_{23}}{C_{33}} - \frac{C_{36}^2}{C_{33}} \right) \partial_1 \partial_2 + \left(C_{26} - \frac{C_{23}C_{36}}{C_{33}} \right) \partial_2^2, \\
(B_u)_{22} &= -\rho_L \partial_t^2 + \left(C_{66} - \frac{C_{36}^2}{C_{33}} \right) \partial_1^2 \\
&\quad + 2\left(C_{26} - \frac{C_{23}C_{36}}{C_{33}} \right) \partial_1 \partial_2 + \left(C_{22} - \frac{C_{23}^2}{C_{33}} \right) \partial_2^2, \\
(B_u)_{33} &= -\rho_L \partial_t^2 .
\end{aligned} \tag{10}$$

By combining Eqs. (6) and (9), we obtain the set of interface conditions as

$$(\boldsymbol{T}^+)_h + h(C\boldsymbol{T}^+)_h = (\boldsymbol{T}^-)_{-h} - h(C\boldsymbol{T}^-)_{-h} + O(h^2), \tag{11}$$

where $\boldsymbol{T}^\pm = (u_1^\pm, u_2^\pm, u_3^\pm, \sigma_{13}^\pm, \sigma_{23}^\pm, \sigma_{33}^\pm)^T$ and the matrix C is given by

$$C = \begin{pmatrix} A_u & A_\sigma \\ B_u & A_u^T \end{pmatrix}. \tag{12}$$

If the thickness of the thin layer, $2h$, is small compared to the wavelengths in the layer, it should be possible to obtain an accurate approximation of the effect of the wave propagation in the layer on the two elastic bodies by replacing the layer by the truncation of Eq. (11),

$$(\boldsymbol{T}^+)_h + h(C\boldsymbol{T}^+)_h = (\boldsymbol{T}^-)_{-h} - h(C\boldsymbol{T}^-)_{-h}. \tag{13}$$

These approximate interface conditions, Eq. (13), are used in the next section to derive an approximate dispersion relation for guided waves in a plate consisting of a thin anisotropic layer sandwiched between two identical isotropic layers.

As would be expected, the interface conditions simplify to the ones derived by Bövik[14] for isotropic materials. As mentioned in the Introduction, the approach taken here is the same as in Ref. 14. If our fields are in the specific form $\boldsymbol{T}^\pm = \boldsymbol{S}^\pm(x_3) e^{i(kx_1 - \omega t)}$, we obtain the interface conditions derived by Rokhlin and Huang.[15] For a good discussion on the physical meaning of the interface conditions, we refer to Ref. 15. Note, however, that the approach used by Rokhlin and Huang is significantly different from Bövik's and ours.

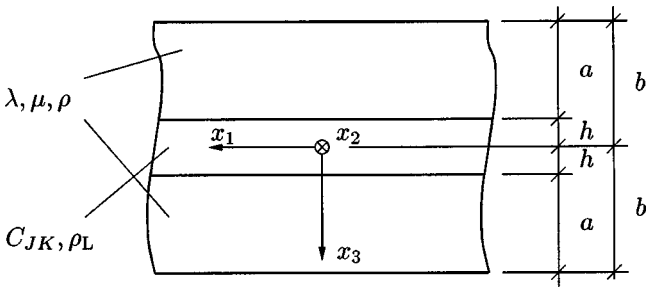


FIG. 2. The anisotropic plate.

Rokhlin and Huang use transfer matrices to describe transmission and reflection by an interface layer. Also, to the best of our knowledge, the above mentioned authors have not used effective interface conditions to study guided waves in plates with a thin interface layer.

B. Dispersion relation

In this section, we use the effective ICs derived in the previous section to obtain an approximate dispersion relation for a layered anisotropic plate. The plate, shown in Fig. 2, consists of a thin anisotropic layer sandwiched between two identical isotropic layers. This configuration is chosen to model a superconducting tape. In the next section, we show some numerical examples for two such systems. Note that the use of the ICs simplifies the analytical treatment of the wave propagation problem significantly since one is now dealing with wave propagation in isotropic layers.

In order to derive the approximate dispersion relation, the thin layer is replaced by the truncated series expansion (13). The equations we then need to solve are

$$\partial_m \sigma_{mj}^\pm = \rho \partial_t^2 u_j^\pm, \quad \pm h \lessgtr x_3 \lessgtr \pm b, \tag{14a}$$

$$\sigma_{jm}^\pm = \lambda \delta_{jm} \partial_n u_n^\pm + \mu(\partial_m u_j^\pm + \partial_j u_m^\pm), \quad \pm h \lessgtr x_3 \lessgtr \pm b, \tag{14b}$$

$$(\boldsymbol{T}^+)_h + h(C\boldsymbol{T}^+)_h = (\boldsymbol{T}^-)_{-h} - h(C\boldsymbol{T}^-)_{-h}, \tag{14c}$$

$$\sigma_{j3}^\pm = 0, \quad x_3 = \pm b, \tag{14d}$$

with the solution in the form $u_j^\pm(x_1, x_2, x_3, t) = v_j^\pm(x_3) e^{i(kx_1 - \omega t)}$.

The general solution to Eqs. (14a) and (14b) may be written in terms of potentials as (see, for example, Achenbach[16]),

$$\begin{aligned}
&\boldsymbol{u}^\pm = \nabla \Phi^\pm + \nabla \times \boldsymbol{\Psi}^\pm, \quad \nabla \cdot \boldsymbol{\Psi}^\pm = 0, \\
&\Phi^\pm = (A_0^\pm \sin p x_3 + B_0^\pm \cos p x_3) e^{i(kx_1 - \omega t)}, \\
&\qquad p = \sqrt{k_p^2 - k^2}, \quad k_p = \omega / c_p, \\
&\Psi_j^\pm = (A_j^\pm \sin q x_3 + B_j^\pm \cos q x_3) e^{i(kx_1 - \omega t)}, \\
&\qquad q = \sqrt{k_s^2 - k^2}, \quad k_s = \omega / c_s,
\end{aligned} \tag{15}$$

where $c_p = \sqrt{(\lambda + 2\mu)/\rho}$ is the longitudinal wave speed in the isotropic material, $c_s = \sqrt{\mu/\rho}$ is the shear wave speed in the isotropic material, and $j = 1,2,3$. The dispersion relation is obtained by applying Eqs. (14c) and (14d) to (15). Due to the symmetry of the problem we split the solution into a

TABLE I. Material properties (C'_{JM} in GPa and ρ in kg/m^3).

Material	C'_{11}	C'_{22}	C'_{33}	C'_{44}	C'_{55}	C'_{66}	C'_{12}	C'_{13}	C'_{23}	ρ
YBCO[a]	268	231	186	37	49	95	132	95	71	6333
Bi-2212[b]	75.8	125.2	125.2	23.15	15.8	15.8	56	56	78.9	6510

[a]Reference 1.
[b]Reference 17.

symmetric and an antisymmetric part. For the symmetric part, the constants in Eq. (15) are related as follows:

$$A_r^+ = -A_r^-, \quad A_s^+ = A_s^-, \quad B_r^+ = B_r^-, \quad B_s^+ = -B_s^-,$$
$$r = 0,3, \quad s = 1,2, \tag{16}$$

and for the antisymmetric part, the constants in Eq. (15) are related as

$$A_r^+ = A_r^-, \quad A_s^+ = -A_s^-, \quad B_r^+ = -B_r^-, \quad B_s^+ = B_s^-,$$
$$r = 0,3, \quad s = 1,2. \tag{17}$$

By replacing the thin layer by the effective ICs, we only need to solve the equations of motion for wave propagation in the isotropic layers. This enables us to explicitly write down the expressions for the symmetric and antisymmetric dispersion relations. For brevity these are omitted. Even though these expressions are significantly more complicated than the ones obtained for guided waves in an isotropic plate, they show the dependence on the material properties of the thin layer explicitly. Interestingly, it is found that the symmetric and antisymmetric modes depend (to this approximation order) on different elastic constants of the thin layer. The symmetric modes only depend on C_{11}, C_{13}, C_{16}, C_{33}, C_{36}, and C_{66}, and the antisymmetric only depend on C_{44}, C_{45}, and C_{55}. For a more thorough discussion on this matter, we refer to Ref. 15. The information about the dependence on the elastic constants should be useful when setting up experiments for material characterization. The accuracy of the approximate dispersion relations is investigated numerically in the next section.

III. NUMERICAL EXAMPLES

In this section we present some numerical examples for guided waves in the three-layered plate shown in Fig. 2. In the dispersion curves, we present the phase velocity $c = \omega/k$ as a function of the frequency $f = \omega/2\pi$, where the pairs (k,ω) are solutions to the exact and approximate dispersion relations. We also show mode shapes for a few selected modes to illustrate the interesting phenomenon of mode interchange.

We consider two different superconducting tapes in the examples. Both tapes have the same dimensions with the thickness of the isotropic layers being $a = 50$ μm and the thickness of the interface layer being $2h = 10$ μm (see Fig. 2 for the definitions of a and h). In the first example, we consider a $YBa_2Cu_3O_{7-\delta}$ (YBCO) layer sandwiched between two nickel (Ni) layers, and in the second, a $Bi_2Sr_2CaCu_2O_{8+\delta}$ (Bi-2212 or BSCCO) layer between two silver (Ag) layers. The properties of the superconducting materials are given in Table I. Note that only the nonzero elastic constants are included. As is seen from the table, YBCO is orthotropic and Bi-2212 is transversely isotropic. The prime on the elastic constants means that they are given in the material's crystal axes system, denoted by (x'_1,x'_2,x'_3). The orientation of the crystal axes system relative to the (x_1,x_2,x_3) system is given by the angle ϕ as shown in Fig. 3. The transformation from the crystal axes system to the latter system is given by Auld.[5] The properties used for silver (Ag) are $\lambda = 81.01$ GPa, $\mu = 30.24$ GPa, and $\rho = 10500$ kg/m^3 and the properties used for nickel (Ni) are $\lambda = 129.5$ GPa, $\mu = 84.71$ GPa, and $\rho = 8910$ kg/m^3.

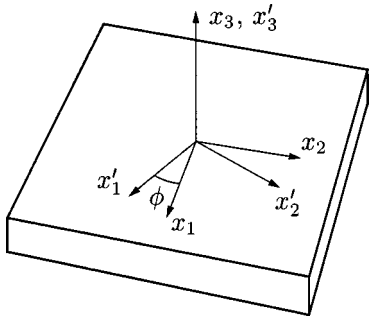


FIG. 3. The orientation of the crystal axes system.

A. The Ni/YBCO/Ni tape

The first tape considered consists of a thin layer made of the orthotropic superconductor YBCO sandwiched between two identical nickel layers. In the figures presented in this section, we consider frequencies in the range 0–100 MHz. In the plots, this means that the wavelengths of the quasi-S waves in the thin layer are no smaller than approximately five times h. The YBCO layer may thus be considered to be relatively thin.

Figure 4 shows a comparison between the layered Ni/YBCO/Ni ($\phi = 0°$) tape and a Ni tape of thickness $2b = 110$ μm. Even though the difference between the curves is not large, it is clearly seen in the figure, already in the S_0 mode. Especially the antisymmetric SH modes are greatly influenced by the YBCO layer. This is probably due to the large difference between C_{44} of YBCO and μ of nickel.

Figures 5 and 6 show comparisons between the dispersion curves obtained by solving the exact dispersion relation and the approximate one derived in Sec. II B. The agreement is excellent up to about $f = 60$ MHz in both figures. At this frequency (60 MHz), the shortest wavelength of the quasi-S waves in the thin layer is approximately eight times h (or four times $2h$). From these figures, the approximation seems to describe the dispersion characteristics of the plate very well in the frequency interval of our interest and should be a very useful analysis tool at lower frequencies.

We end this section by considering the characteristics of the dispersion curves in greater detail. In Fig. 7, a magnification of the dispersion curves (SH_0 and S_0 if $\phi = 0°$) is shown for different directions of propagation (different values of ϕ). The notable thing in this figure is the fact that the dispersion curves cannot cross in the off-angle cases ($\phi \neq n\pi/2$, $n = 0,1,\ldots$), the reason being that the in-plane and out-of-plane displacements couple. Therefore, mode inter-

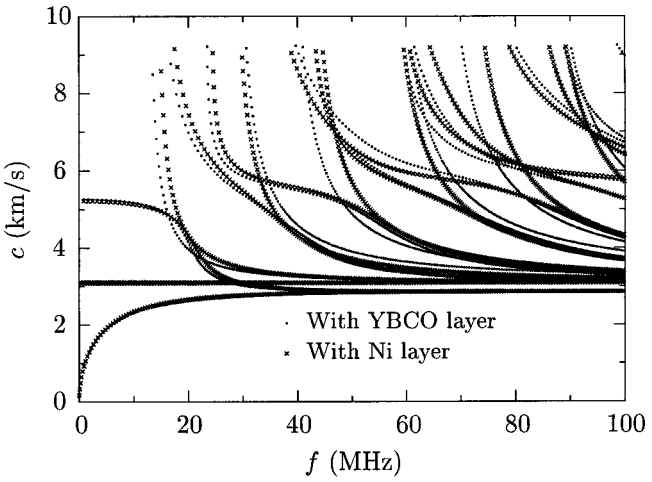


FIG. 4. Dispersion curves with an YBCO layer ($\phi=0°$) and a Ni layer.

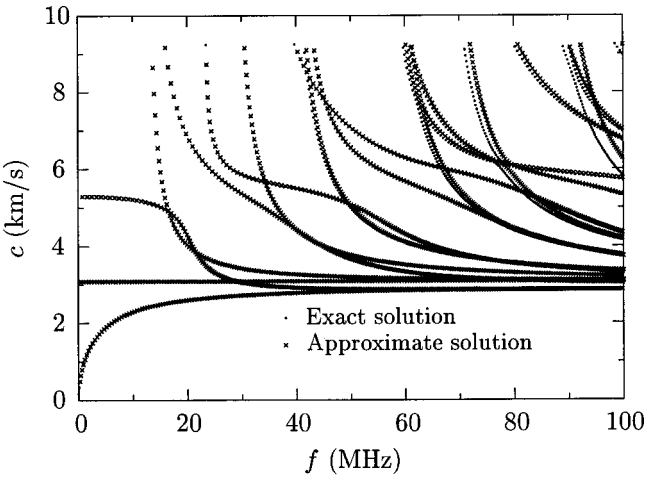


FIG. 6. The exact versus the approximate curves for Ni/YBCO/Ni ($\phi=45°$).

change will take place, and in a region close to the point where the curves are at the smallest distance from each other, a strong mixing of the modes should occur. We also note that the location of this point and the size of the gap between the curves seem to be fairly sensitive to the anisotropy and the thickness of the layer.

In Fig. 8, we show the normalized absolute values of the components of the mode shapes on the surface of the tape ($x_3=-b$) as a function of the frequency for the two different modes in Fig. 7 when $\phi=45°$. Here, the mode interchange and mixing of the modes are clearly seen. The slower mode (quasi-*SH* $\rightarrow$ quasi-longitudinal) in Fig. 7 goes from being mostly polarized in the x_2 direction into being mostly polarized in the x_1x_3 plane. The opposite should obviously hold for the faster mode (quasi-longitudinal $\rightarrow$ quasi-*SH*), and this is also seen in the figure. A very interesting thing seen in Fig. 8 is that all the components of the mode shapes are of a significant size in a fairly wide interval centered around the point where the modes are closest to each other. This could prove to be useful in ultrasonic measurements since all three displacement components could be used to excite and detect these modes.

B. The Ag/Bi-2212/Ag tape

The second system considered consists of a thin interface layer made of the transversely isotropic superconductor Bi-2212 and two identical isotropic layers made of silver. In this case, we consider the frequency range 0–50 MHz. In the plots, the wavelengths of the quasi-*S* waves in the Bi-2212 layer are no smaller than approximately six times *h*. This means that the Bi-2212 layer may be considered to be relatively thin.

We compare the effect of the introduction of a different layering material on the dispersion characteristics in Fig. 9. Here, we show a comparison between the exact dispersion curves when the layer is made of Bi-2212 ($\phi=0°$) and when it is made of silver (resulting in a pure Ag plate of thickness $2b=110$ μm). For this material combination, the difference between the curves is fairly small.

Figure 10 shows a comparison between the exact dispersion curves and the ones obtained from the approximation derived in Sec. II B. Here, we find that the approximation works very well up to about 35 MHz. At this frequency, the shortest wavelength of the quasi-*S* waves in the Bi-2212 layer is approximately nine times *h*.

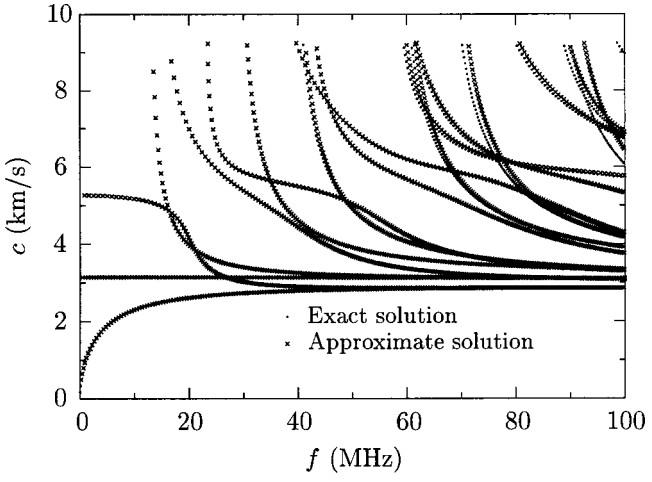


FIG. 5. The exact versus the approximate curves for Ni/YBCO/Ni ($\phi=0°$).

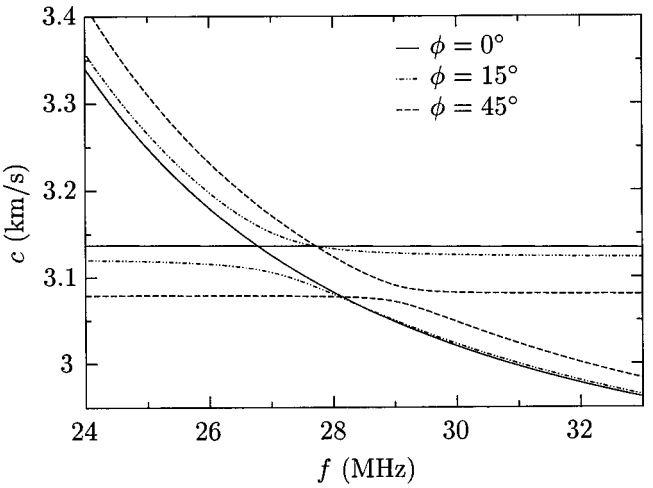


FIG. 7. A magnification of the dispersion curves for Ni/YBCO/Ni.

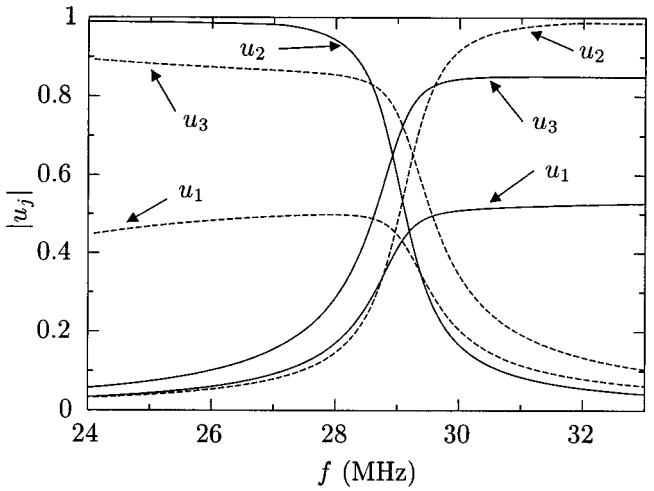


FIG. 8. Mode shapes on the surface of the plate corresponding to Fig. 7 for $\phi=45°$. The solid lines give the slower mode, the dashed give the faster mode.

The interchange of modes for the off-axes propagation is observed in this case also. The first such interchange (between S_0 and SH_0) occurs now at a lower frequency.

IV. CONCLUDING REMARKS

In summary, we present the derivation of the simplified interface conditions in a three-layered plate made of a *thin* anisotropic interface layer sandwiched between two identical isotropic layers. The approximation is obtained by expanding the displacement and traction through the thickness of the thin layer and from this, obtaining approximate interface conditions. The use of the approximate interface conditions enables the derivation of a closed form (approximate) dispersion relation for the three-layered system.

In the numerical examples, two systems, each consisting of a thin superconducting layer sandwiched between two identical isotropic layers, are considered. The first system consists of an YBCO layer sandwiched between two Ni layers and the second consists of a Bi-2212 layer sandwiched between two Ag layers. Comparisons between plates with and without the superconducting interface layers are made.

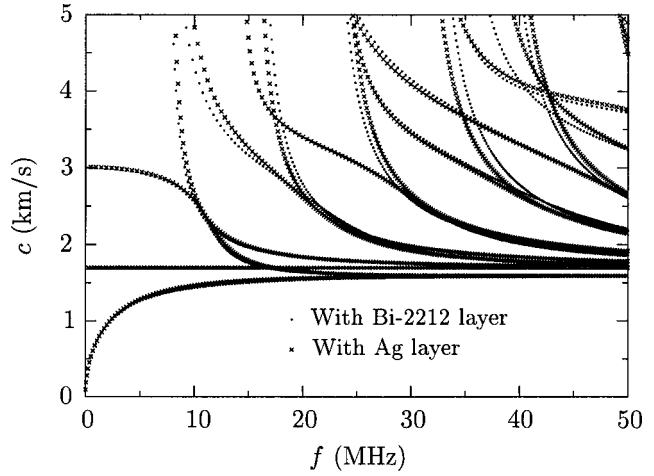


FIG. 9. Dispersion curves with a Bi-2212 ($\phi=0°$) layer and an Ag layer.

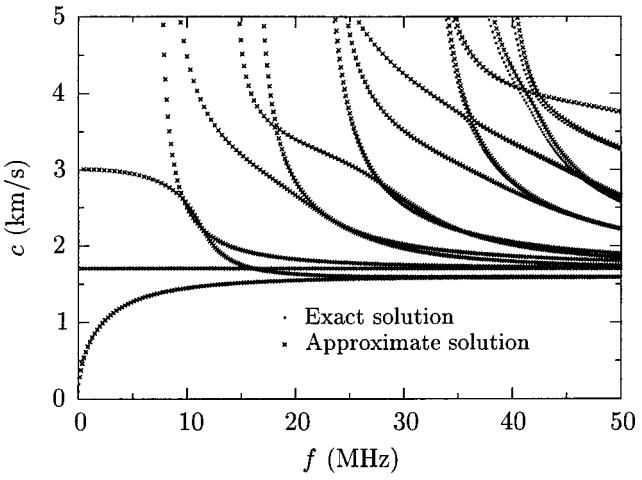


FIG. 10. The exact versus the approximate curves for Ag/Bi-2212/Ag ($\phi=45°$).

We compare the superconducting plates with plates having an isotropic interface layer made of the same material as the outer layers. The effect of the presence of the superconductors is clearly seen (even at fairly low frequencies). Comparisons between exact and approximate dispersion curves are made and it is found that the approximation works very well in the frequency intervals of our interest. The approximation should therefore be a very useful tool in analyzing the dependence and sensitivity of the curves on the material properties of the thin layer. Finally, the dispersion characteristics of the layered plates are studied in greater detail. Due to the anisotropy of the superconducting layers, in-plane and out-of-plane modes will couple (symmetric and antisymmetric modes will still decouple). This means that certain modes cannot cross in directions that do not coincide with the anisotropic material's symmetry planes. This effect is shown in a magnification of the dispersion curves for the Ni/YBCO/Ni tape for different directions of propagation. It is seen that the location of the point where the modes are at the shortest distance from each other and the size of this gap are fairly sensitive to the direction due to the anisotropy. This is very encouraging, since one of the main objectives of this study is to investigate if useful data about the anisotropy may be extracted from ultrasonic experiments at relatively low frequencies. In addition, the components of the mode shapes on the surface of the plate are shown as functions of the frequency for the magnified part of the dispersion curves. Here, the mode interchange is clearly seen and more importantly, the interval where strong mixing occurs is fairly wide. In this interval, all components of the displacement are of a significant size. This is very interesting since this could allow excitement and detection via any of the displacement components.

ACKNOWLEDGMENTS

The work of the first author (A.J.N.) has been financially supported by the Swedish Foundation for International Cooperation in Research and Higher Education (STINT) and this is gratefully acknowledged. The work of S.K.D. and M. L.D. was supported by a grant from the Engineering Re-

search Program, Office of Basic Energy Sciences, DOE (Grant No. DE-FG03-97ER14738). S.K.D. also acknowledges the Faculty Fellowship support from the University of Colorado, Boulder, which enabled him to spend some time at the Department of Mechanics at Chalmers University of Technology, and the hospitality he received there.

[1] M. Lei, J. L. Sarrao, W. M. Visscher, T. M. Bell, J. D. Thompson, A. Migliori, U. W. Welp, and B. W. Veal, ''Elastic constants of a monocrystal of superconducting $YBa_2Cu_3O_{7-\delta}$,'' Phys. Rev. B **47**, 6154–6156 (1993).

[2] Y. Shindo, H. Ledbetter, and H. Nozaki, ''Elastic constants and microcracks in $YBa_2Cu_3O_7$,'' J. Mater. Res. **10**, 7–10 (1995).

[3] *Physical Acoustics*, edited by M. Levy (Academic, New York, 1993), Vol. 20.

[4] G. W. Farnell and E. L. Adler, ''Elastic wave propagation in thin layers,'' in *Physical Acoustics*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1972), Vol. 9, Chap. 2, pp. 35–127.

[5] B. A. Auld, *Acoustic Fields and Waves in Solids* (Krieger, Malabar, FL, 1990), Vols. 1 and 2.

[6] A. H. Nayfeh, *Wave Propagation in Layered Anisotropic Media with Applications to Composites* (Elsevier, Amsterdam, 1995).

[7] Y. Li and R. B. Thompson, ''Influence of anisotropy on the dispersion characteristics of guided ultrasonic plate modes,'' J. Acoust. Soc. Am. **87**, 1911–1931 (1990).

[8] L. P. Solie and B. A. Auld, ''Elastic waves in free anisotropic plates,'' J. Acoust. Soc. Am. **54**, 50–65 (1973).

[9] W. Karunasena, A. H. Shah, and S. K. Datta, ''Wave propagation in a multilayered laminated cross-ply composite plate,'' ASME J. Appl. Mech. **58**, 1028–1032 (1991).

[10] D. E. Chimenti, ''Guided waves in plates and their use in materials characterization,'' Appl. Mech. Rev. **50**, 247–284 (1997).

[11] S. K. Datta, ''Wave propagation in composite plates and shells,'' in *Comprehensive Composite Materials*, edited by T.-W. Chou (Elsevier, Oxford, 2000), Vol. 1, Chap. 18, pp. 511–558.

[12] J. L. Rose, *Ultrasonic Waves in Solid Media* (Cambridge University Press, Cambridge, 1999).

[13] A. J. Niklasson, S. K. Datta, and M. L. Dunn, ''On approximating guided waves in plates with thin anisotropic coatings by means of effective boundary conditions,'' J. Acoust. Soc. Am. **108**, 924–933 (2000).

[14] P. Bövik, ''On the modelling of thin interface layers in elastic and acoustic scattering problems,'' Q. J. Mech. Appl. Math. **47**, 17–42 (1994).

[15] S. I. Rokhlin and W. Huang, ''Ultrasonic wave interaction with a thin anisotropic layer between two anisotropic solids. II. Second-order asymptotic boundary conditions,'' J. Acoust. Soc. Am. **94**, 3405–3420 (1993).

[16] J. D. Achenbach, *Wave Propagation in Elastic Solids* (North-Holland, Amsterdam, 1973).

[17] M. Boekholt, J. V. Harzer, B. Hillebrands, and G. Güntherodt, ''Determination of the sound velocities and the complete set of elastic constants for $Bi_2Sr_2CaCu_2O_{8+\delta}$ single crystal using Brillouin light scattering,'' Physica C **179**, 101–106 (1991).

Ultrasonic cavitation monitoring by acoustic noise power measurement

J. Frohly,[a] S. Labouret, C. Bruneel, I. Looten-Baquet, and R. Torguet
Institut d'Electronique et de Microélectronique du Nord, Département OAE-U.M.R. CNRS 9929, ENSIMEV, Université de Valenciennes et du Hainaut-Cambrésis, BP 311-59304 Valenciennes Cedex, France

(Received 10 June 1996; revised 15 June 2000; accepted 28 July 2000)

In this paper, a new tool is proposed to carry out acoustic cavitation monitoring and to have an overview of its effects in applications. After a brief review of the cavitation characterization techniques, it is shown that cavitation noise is a suitable and accurate indicator of the cavitation activity induced in a liquid. In the first part of this study, the origin of the first spectral component of the cavitation noise is discussed. The $f/2$ and $2f$ component evolution measurement at a driving frequency around 1 MHz confirms Neppiras' ones and gives an indicator of the cavitation inception. In the second part, the cavitation noise spectrum distortion is considered as a function of the acoustic power transmitted to the liquid in order to obtain an indicator of cavitation activity. In the last part, this new tool is used to bring to the fore the hysteresis effect associated with the cavitation. An experimental correlation between cavitation noise power measurement and the sonochemical activity in an oxidization process is also presented.
[S0001-4966(00)01211-X]

PACS numbers: 43.25.Yw [MAB]

LIST OF SYMBOLS

R_0	initial bubble radius
R	bubble radius at any given moment
$\dot{R}$	bubble wall velocity
$\ddot{R}$	bubble wall acceleration
P_A	maximum amplitude of the acoustic pressure
P_0	equilibrium pressure in the liquid
σ	surface tension
f	driving frequency
ρ_0	density of the liquid
ω	angular frequency $=2\pi f$
n	gas polytropic index
f/q	with q integer: subharmonic of the driving frequency
pf	with p integer: harmonic of the driving frequency
pf/q	with p and q integers: ultraharmonic of the driving frequency

I. INTRODUCTION

Cavitation is defined as the formation of bubbles in a liquid subject to local pressure variations. Thus, the propagation of an acoustic wave through a liquid may induce cavitation. The effects associated with cavitation onset in a liquid cover a wide variety of physical phenomena such as luminescence, shock wave emissions, and free-radicals formation. In the past 20 years, most of the research carried out on acoustic cavitation shows that this phenomenon should be a powerful tool for many industrial applications. Cleaning systems using cavitation are now well-known. Sonochemistry, a process in which cavitation is used as a kind of chemical catalyst, has been studied for 50 years. Moreover, some authors[1–3] have also reported interesting results concerning the bacteriological effects of cavitation. Despite these encouraging prospects, ultrasonic cavitation still remains a laboratory tool for two main reasons.

Uncertainties concerning the mechanisms whereby cavitation enhances chemical reactivity or generates light emissions are the first reasons for limited industrial scale development of cavitation. Many works have been carried out in order to explain the behavior of a cavitation bubble. Sehgal *et al.*,[4] and later Suslick[5–7] showed that collapsing bubbles may generate hot spots in the liquid. Thus, they concluded that sonochemistry and sonoluminescence arise from localized peak temperatures and pressures. Recently, Lepoint[8,9] developed an electrical interpretation of these phenomena. He assumed that the electrical charge carried by a bubble might play a pivotal part in sonochemical mechanisms. Localized electrical fields produced by bubble distortions may generate microscopic intracavity discharges during the collapse. He concludes that the microplasma induced in the bubble is the source of sonochemical and sonoluminescent activity. The sonoluminescence emission spectrum analysis, currently receiving a great deal of attention, should rapidly provide more information about cavitation mechanisms.

The lack of cavitation bubble field probes is the second reason that limits the industrial-scale development of cavitation. An obvious reason for the complexity of cavitation characterization is the great disparity in bubble characteristics inside the cavitation field. Once it is generated, an acoustic cavitation bubble may undergo two quite different kinds of radial oscillations. It may oscillate nonlinearly during many cycles of the acoustic wave, termed ''stable cavita-

[a]Electronic mail: frohly@univ-valenciennes.fr

tion,'' or it may grow rapidly and collapse more or less violently, termed ''inertial or transient cavitation.'' Neppiras[10] and Apfel[11,12] showed that the bubble behavior depends both on the bubble size compared to the acoustic wavelength and on the acoustic wave amplitude. Despite its fundamental interest, the estimation of these two parameters is a real problem in cavitation studies, even if acoustic amplitude may be determined with calibrated transducers. In fact, the cavitation state induced in the liquid is rarely described in most experimental works involving cavitation. So, easy monitoring methods have to be developed to correlate the cavitation state produced in a liquid to the chemical or biological effects observed. Some experimental methods for characterizing the cavitation bubble fields are described in the following paragraphs.

The easiest way to study cavitation may seem to be the direct observation of the bubble field. Because of the microscopic size and the rapid oscillations of each cavitation bubble, the direct observation of the phenomenon requires a very high spatial and temporal resolution photographic device. Lauterborn[13–15] and his co-workers provided an important contribution to cavitation bubble studies by using high-speed photography. They also showed that holocinematography associated with digital processing of the three-dimensional images allows the cavitation bubble fields to be monitored. Unfortunately, the experimental device is too difficult to implement in industrial applications.

As cavitation bubble field visualization is quite complex, workers have used indirect observations of macroscopic criteria in order to describe the cavitation state produced in a liquid. The criterion selected to monitor for cavitation in industrial applications must combine real-time measurements on wide size ranges with easy installation and setup. We have considered most of the existing experimental techniques for cavitation detection. All of these techniques are based on the analysis of the interaction between the wave and the bubbly liquid. The optical solution, which consists of the analysis of the diffraction pattern obtained when a laser light propagates through a bubbly liquid, allows reliable, accurate, and nonintrusive bubble sizing.[16] Unfortunately, we noticed that the sound wave which induces cavitation produces an acousto-optic effect. Thus, the diffraction pattern recorded results from both the interaction between the laser beam and the bubbles and the acousto-optic effect. It is therefore quite difficult to obtain information concerning an acoustic cavitation bubble field in this way. As it has been well-established that the presence of air bubbles in water can change the velocity and attenuation of an acoustic wave, we have considered an acoustical solution to characterize cavitation. Van Wijngaarden,[17] and later Caflish[18] and Miksis,[19] have modeled the transmission and scattering of an acoustic wave through a bubbly liquid. Moreover, it has been shown that the measurement of attenuation[20–23] and velocity[24] of a low-power acoustic pulse propagating through a bubbly liquid provides quantitative information about the bubble field. However, this approach cannot be applied to the acoustic cavitation bubble field characterization because it assumes linear pulsations of stable bubbles about an equilibrium radius. Another acoustic solution to study cavitation has been proposed by Roy[25,26] and co-workers. The authors used a backscattering technique for the detection of transient cavitation. Signal processing applied to the backscattered wave could probably be used to monitor cavitation, but this method is very difficult to implement in industrial applications because of focusing and size constraints.

This list of methods has been limited to the easiest techniques. However, the discussion of their application to cavitation characterization for industrial applications shows that appropriate tools are needed to monitor cavitation. From this viewpoint, we have developed a cavitation probe based on the cavitation noise measurement. After a rapid review of different spectrum contributions present in the acoustic cavitation spectrum, we show that the spectral distribution of the cavitation noise is strongly linked to the cavitation state. Then, we propose the cavitation noise power, obtained by numerical integration of the spectra, as a new criterion to represent the cavitation state. We show that this criterion may be used to bring to the fore the hysteresis effect associated with cavitation phenomenon. Finally, we present the experimental correlation between this criterion and the sonochemical reactivity in an oxidization process.

II. BUBBLE OSCILLATIONS IN AN ACOUSTIC FIELD

In this study, we limit our interest to the case of a small bubble compared to the acoustic wavelength. Much work has been done concerning the dynamics of a single bubble subject to an acoustic wave. Noltingk and Neppiras[27] developed the first dynamic model of bubble oscillations. More complete models for the description of the dynamics of a spherical bubble have been proposed by Kirkwood and Bethe,[28] Gilmore,[29] and later by Prosperetti.[30] All of these models lead to a complex and nonlinear equation requiring a numerical solution. Prosperetti[31] and Lauterborn[32] have brought about a wide contribution to these computations. They have shown that the oscillating behavior of a bubble, which is strongly dependent on the acoustic wave amplitude and also on the relationship between the bubble size and the acoustic frequency, may be characterized by the spectrum of the radius–time function.

Whatever the model considered, the numerical computation of the radius–time variation shows that the frequency distribution of the oscillating motion may contain harmonics, subharmonics, and ultraharmonics of the driving frequency. The harmonic components arise from the strongly nonlinear motion of the bubble subject to forced oscillations. In the next two sections, we discuss the situations leading to subharmonic and/or ultraharmonic oscillations of the bubble. These occurrences differ essentially in terms of the acoustic wave intensity.

A. Subharmonic oscillations at low intensities

At low intensities, subharmonic oscillations result from bubbles that are twice the resonance size. These bubbles may present parametric oscillations for weak variations of their radius R around the initial value R_0.

This phenomenon, already mentioned by Flynn[33] and Mei,[34] may be described by a simple analytical model. Con-

sider the Rayleigh–Plesset–Noltingk and Neppiras equation

$$R\ddot{R}+\frac{3}{2}\dot{R}^2+\frac{1}{\rho_0}\left[P_0-P_A\sin\omega t+\frac{2\sigma}{R}-\left(P_0+\frac{2\sigma}{R_0}\right)\left(\frac{R_0}{R}\right)^{3n}\right]=0. \quad (1)$$

For small amplitude radial oscillations, we may write

$$R=R_0(1+x), \quad (2)$$

where $x\ll 1$ and has the following form: $x=a\exp(j\omega t)$, where $a\ll 1$. So,

$$\dot{R}=R_0\dot{x};\quad \ddot{R}=R_0\ddot{x};\quad \frac{1}{R}\cong R_0(1-x);\quad \left(\frac{R_0}{R}\right)^{3n}\cong 1-3nx.$$

Thus, motion equation (1) becomes

$$R_0^2\ddot{x}+R_0^2x\ddot{x}+\frac{3}{2}R_0^2\dot{x}^2+\frac{1}{\rho_0}\left[P_0-P_A\sin\omega t+\frac{2\sigma}{R_0}(1-x)-\left(P_0+\frac{2\sigma}{R_0}\right)(1-3nx)\right]=0. \quad (3)$$

Equation (3) may also be written

$$\ddot{x}+x\ddot{x}+\frac{3}{2}\dot{x}^2+\frac{1}{\rho_0R_0^2}\left[-P_A\sin\omega t+x\left\{3n\left(P_0+\frac{2\sigma}{R_0}\right)-\frac{2\sigma}{R_0}\right\}\right]=0. \quad (4)$$

Writing

$$\Omega_0^2=\left[3n\left(P_0+\frac{2\sigma}{R_0}\right)-\frac{2\sigma}{R_0}\right]\frac{1}{\rho_0R_0^2}, \quad (5)$$

and

$$\Omega_0^2h=\frac{P_A}{\rho_0R_0^2}, \quad (6)$$

Eq. (4) is then written

$$\ddot{x}+x\ddot{x}+\tfrac{3}{2}\dot{x}^2+\Omega_0^2x=\Omega_0^2h\sin\omega t. \quad (7)$$

Equation (5) reduces to Minnaert's expression[35] for small σ and $n=\gamma$.

Neglecting the $\dot{x}^2$ term and dividing by $(1+x)$ factor with the approximation $1/(1+x)\cong 1-x$, Eq. (7) becomes

$$\ddot{x}+\Omega_0^2x=\Omega_0^2h(1-x)\sin\omega t, \quad (8)$$

or

$$\ddot{x}+\Omega_0^2(1+h\sin\omega t)x=\Omega_0^2h\sin\omega t. \quad (9)$$

The first member of Eq. (9) corresponds to the equation of a parametric oscillator described by Landau and Lifchitz.[36]

This fact explains that in an acoustic cavitation field, bubbles having a R_0 radius corresponding to a resonance pulsation $\Omega_0=\omega/2$, with ω the driving pulsation, will present a strong oscillating component at the $f/2$ subharmonic.

The parametric oscillation amplitude increases exponentially in function of time with an amplification coefficient $s=h\Omega_0/4$ given by Landau and Lifchitz[36] in the precise case of the half harmonic.

In spite of the drastic approximations made in the previous calculation, we may verify that an expression in the following form $ae^{+st}\cos(\Omega_0t+\varphi)$ is a solution of expression (7) directly derived from the Rayleigh–Plesset–Noltingk–Neppiras equation (1) for low intensities.

Introduction of such a solution in relation (7) gives an s value equal to $s=h\Omega_0/12$. This amplification coefficient is lower than the previous one due to the damping effect ($\dot{x}^2$ term) taken into account in the latter case.

In conclusion, this result shows that the subharmonic oscillation appears even at low intensities in a randomly distributed cavitation bubble field, when bubbles are twice the resonance size for a driving frequency f. So, this component cannot be considered a significant parameter of a particular activity of cavitation bubbles.

B. Subharmonic and ultraharmonic oscillations at high intensities

Using methods of chaos physics to study the problem, Lauterborn[37] has described stochastic oscillations of a bubble. In this case, the spectral distribution of the radius–time function, obtained by Fourier transform, contains both subharmonic and ultraharmonic components. Each bubble of the cavitation -field may present chaotic oscillations for sufficiently high amplitude of the acoustic excitation. The higher the amplitude of the acoustic wave, the more ultraharmonic oscillations occur.

C. Cavitation noise induced by bubble oscillations

As chemical or biological effects induced by a bubble on the liquid are likely to be strongly linked to its oscillating behavior, it is paramount to evaluate this parameter. From an industrial viewpoint of cavitation applications, this problem is of fundamental interest. Spectral analysis of the cavitation noise should be a solution. In fact, each bubble of the cavitation field behaves as a secondary source of acoustic emission. So, the characteristics of each secondary source depends on the bubble motion. Llyichev[38,39] has characterized the pressure wave emitted by a bubble as a function of the driving pressure. He has concluded that the spectrum of this pressure wave may contain harmonics, subharmonics, and ultraharmonics of the driving frequency as well as the oscillating motion of the bubble. These theoretical results confirm the experimental observations achieved by Neppiras.[10]

In the following parts, we will show that the cavitation noise spectrum analysis is efficient to characterize the cavitation activity. For example, the previous considerations have shown that the half-subharmonic emission may arise from two very different phenomena. So, this criterion, currently used as a cavitation probe, is ambiguous and a lot of work is still needed to distinguish the contribution of parametric oscillations of twice the resonant size bubbles from the contribution of chaotic oscillations to the subharmonic emissions. Because of the exponential increase in oscillation amplitude, the lifetime of twice the resonant size bubbles is very short in the liquid and their subharmonic emission at $f/2$ probably consists of short bursts. Conversely, the subharmonic and ultraharmonic emissions due to chaotic oscilla-

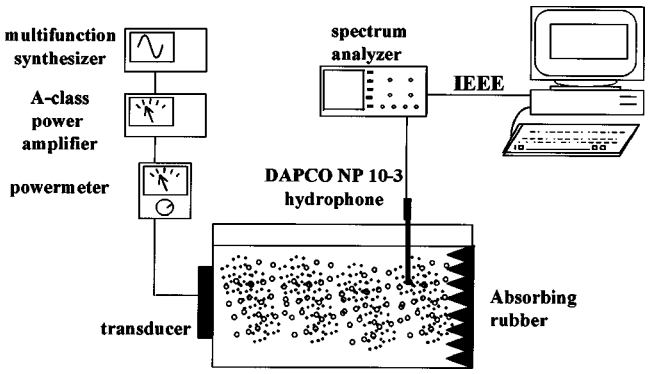


FIG. 1. Experimental setup.

tions of bubbles seem to be stable in time. This report should be a starting point for further investigations. In the meantime, in this paper we propose a new tool to characterize both the cavitation state and the cavitation activity in the liquid.

III. EXPERIMENTAL SETUP AND RESULTS

As a first step, we have observed the cavitation noise spectrum distortion as a function of the acoustic intensity. This value is deduced from the measured electrical power, knowing the electromechanical conversion factor of the transducer determined by a back-reflection technique.[40] The experimental setup is presented in Fig. 1. The liquid used is tap water saturated with air, at 20 °C and at atmospheric pressure. The transducer consists of a 6×6-cm piezoelectric transducer (PZT) ceramic bonded to a stainless-steel substrate. It is driven at 1.075 MHz. The cavitation noise is recorded with a broadband hydrophone DAPCO NP 10-3, and this signal is transmitted to an analogical spectrum analyzer, digitized, and stored in a computer. Cavitation noise spectrum is expressed in a logarithmic scale.

The results are presented in Fig. 2. These observations are quite similar to Esche.[41] The cavitation noise spectrum at low acoustic intensities, typically under 24 mW/cm^2, is a line spectrum containing a fundamental line and harmonic, subharmonic, and ultraharmonic components. At high intensities, a broadband component is superimposed on this line spectrum. This signal, called ''white noise'' by Neppiras,[10] may be attributed to two phenomena: strong chaotic oscillations of bubbles and shock waves emission by collapsing bubbles.

Llyichev[38,39] has shown that ultraharmonic contributions from the bubble's oscillating motion increase considerably at high intensities. Lauterborn[32] has shown that all the spectral components broaden when intensity rises. Furthermore, Cordry[42] and Crum have proposed another possibility, the emission of a shock wave by a single bubble. During the collapse, the bubble wall velocity is so high that a shock wave is released in the liquid. The shock wave produces a broadband contribution to the cavitation noise. The large amount of theoretical studies, as well as the observations of the motion of a single bubble in a standing wave field by Gaitan *et al.*[43] have greatly contributed to an understanding of the oscillating behavior of each bubble. Then, it is possible to associate a particular motion of the bubble with each spectral component of the cavitation noise.

In Figs. 3 and 4, we show the $f/2$ and $2f$ spectrum component evolution measurements for different acoustic power. We can observe a sudden increase in these components for low intensities up to 60 mW/cm^2 in our experimental conditions. These results are in agreement with these of Neppiras.[10] For higher acoustic power values, a threshold appears. These $f/2$ and $2f$ peaks show parametric or nonlinear bubble oscillations to be found in stable cavitation. They do not represent by themselves an indicator of transient cavitation state. In fact, $f/2$ and $2f$ peak levels reach a quasi-saturation value from 60 mW/cm^2 acoustic intensity, while we can see in Fig. 2 that noise contribution continues to increase significantly up to 120 mW/cm^2. The transient cavitation and bubble collapses produce white noise. The qualitative observation of the spectra shows the emergence and the increase in the cavitation noise with the acoustic intensity. So, it would be interesting to have a quantitative indicator of the transient cavitation correlated to the bubble collapses and to the cavitation bubble activity.

There are two possibilities to assess this overall contribution. The first way would consist of measuring a significant physical parameter by a cavitation spectrum integration in a linear scale. Then, we should obtain an estimation of the real cavitation noise power value. But, in this case, the anharmonic noise weakly contributes to the final result. The second way consists of a cavitation spectrum integration in a logarithmic scale in order to enhance the white-noise contribution. Although this integration does not really correspond to a physical power we will consider it, in the following, as a cavitation noise power indicator called CNP. The resulting data are linked to the cavitation noise power[44] within a frequency span corresponding to the hydrophone bandwidth. The result is presented in Fig. 5. This curve confirms the qualitative observations and the 120-mW/cm^2 characteristic value previously mentioned. We also observe that the cavitation noise power reaches a maximum value P_L. In Fig. 6 we present the respective noise and peak contributions to the CNP obtained by cutting the acoustic spectrum at a 30-dB level value corresponding to the maximum value of the white noise (see Fig. 2). The part over 30 dB has been considered as the peak contribution and the part below this value has been considered as the noise contribution. We may conclude from the comparison of both parts that the CNP is an appropriate indicator of the white noise since the subharmonic, harmonic, and ultraharmonic components contribute less than 10% to the CNP value.

In conclusion we may consider that the $f/2$ and $2f$ peak levels are good tools for the setting up of stable cavitation regime while the CNP is a good indicator to characterize transient cavitation state. The respective acoustic intensity thresholds are, in our case, around 60 and 120 mW/cm^2. Furthermore we have noticed a maximum value of CNP. In the next parts of this paper, we will normalize the cavitation noise power measurements in comparison with the limit value P_L. Using this new probe, we have studied two phenomena induced by acoustic cavitation in a liquid: a hyster-

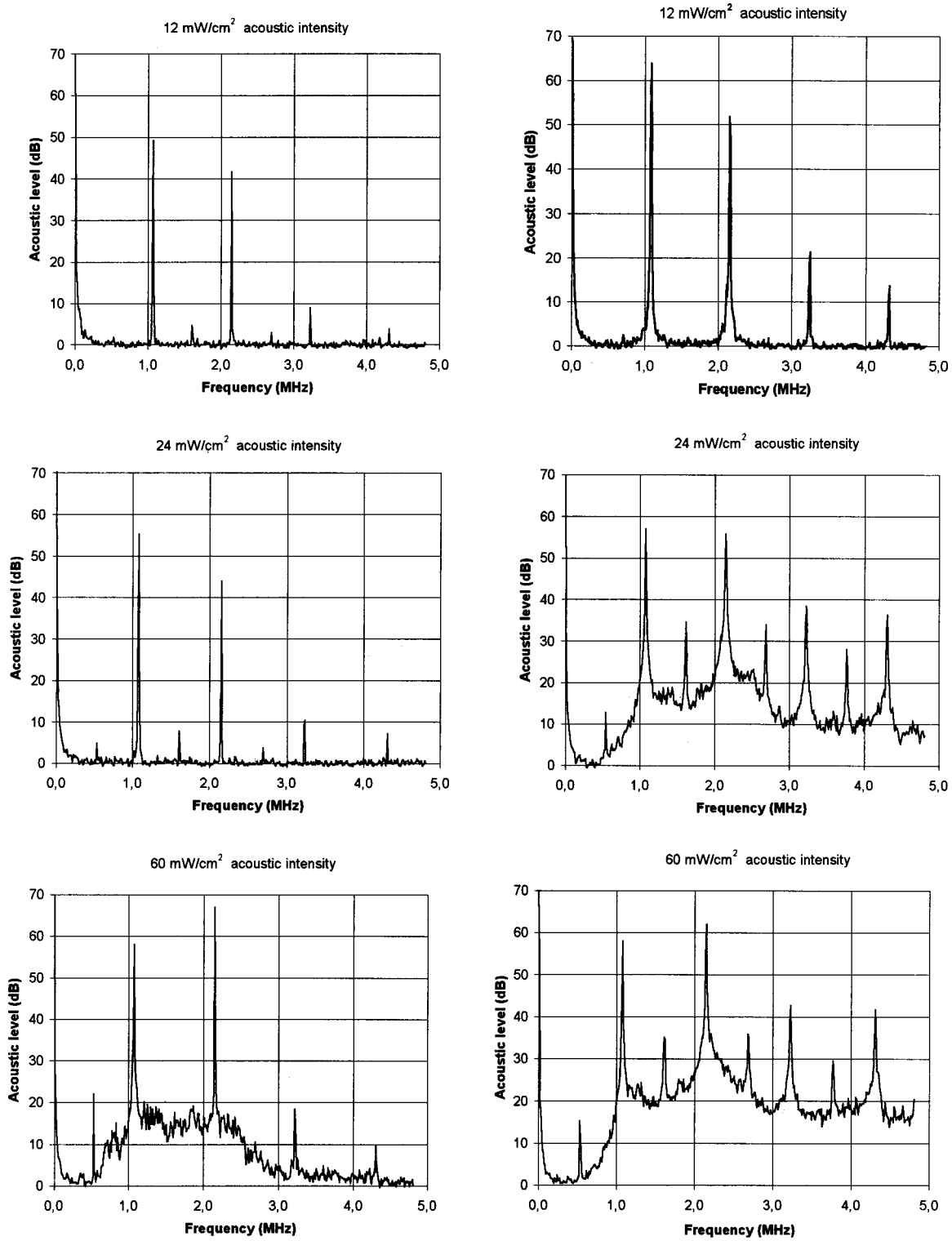


FIG. 2. Cavitation noise spectrum distortion as a function of the acoustic intensity transmitted to the liquid, for increasing power (left curves) and decreasing power (right curves).

esis effect linked to the increase in cavitation nuclei stabilized in the liquid and the sonolysis of water induced by free-radicals formation.

IV. HYSTERESIS EFFECT

Neppiras[10] has shown that the cavitation state induced in a liquid depends on the sample history. Our observations of $f/2$ and $2f$ spectrum component levels do not permit us to conclude that there is a significant hysteresis effect (Figs. 3 and 4). But, Fig. 2 clearly indicates a huge difference in acoustic spectra with regards to their history. We can also observe that spectral components broaden when acoustic power increases. So, we confirm the Lauterborn results. But, spectral component widening is maintained when acoustic power decreases. In order to detect the hysteresis phenom-

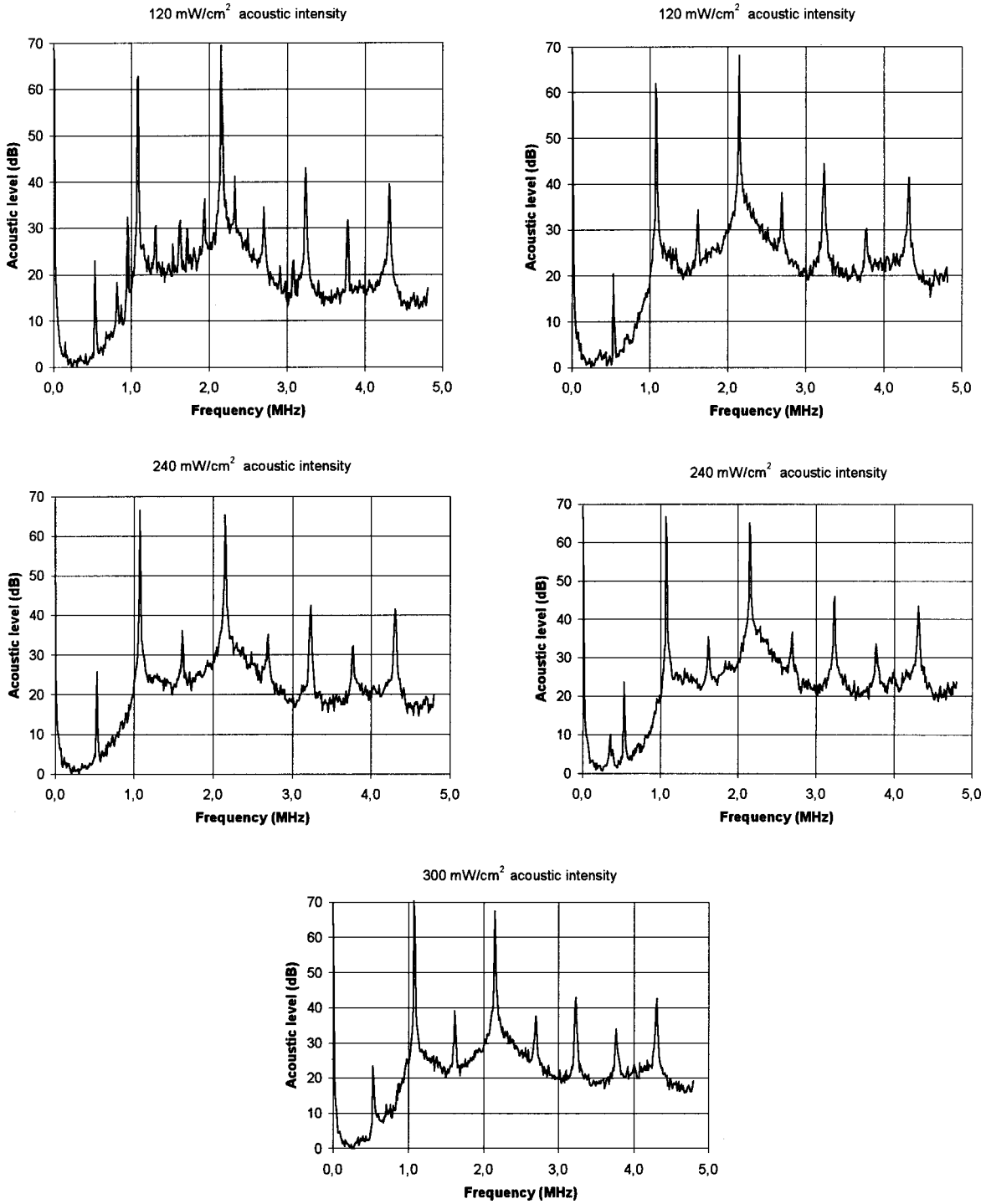


FIG. 2. (*Continued.*)

enon, we have recorded the cavitation noise power during an increasing power sweep, and immediately after a decreasing power sweep. The result obtained at a 1.075-MHz frequency is presented in Fig. 7. This curve clearly confirms the hysteresis effect, and shows that once generated in the liquid, a cavitation state needs a lower driving power to be maintained. This result might present a real industrial interest when this phenomenon takes place in the case of a large liquid volume locally insonified. We will check this point later.

V. SONOLYSIS STUDY

Cavitation bubble collapse induces free-radical generation in the liquid. Chemical effects of cavitation arise from this phenomenon. As the lifetime of these chemical species is very short, it is therefore very difficult to measure their concentration directly. So, their formation is brought to the fore using an oxidization process. We have established an experimental correlation between the cavitation noise power and the chemical reactivity in the Fe^{2+}/Fe^{3+} oxidization reaction.[45] The driving frequency is 1.6 MHz. The cavitation

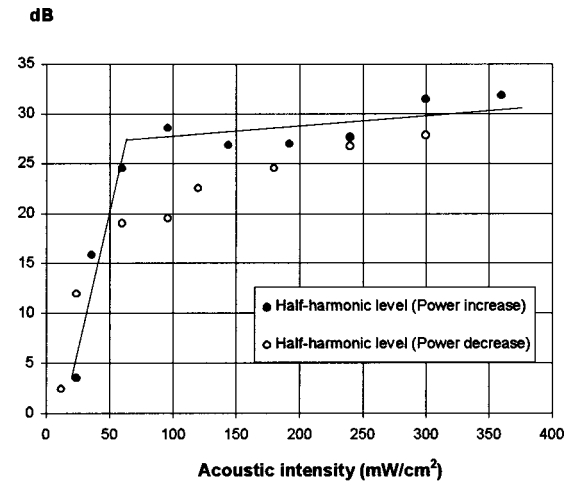


FIG. 3. Half-harmonic evolution as a function of the acoustic intensity transmitted to the liquid.

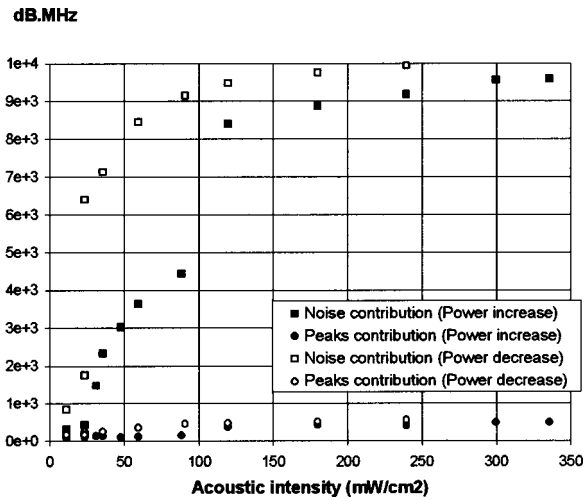


FIG. 6. Comparison between noise and peak contributions to the cavitation noise power.

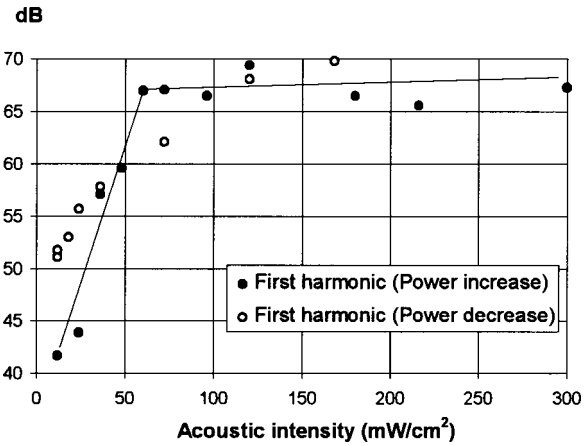


FIG. 4. First harmonic evolution as a function of the acoustic intensity transmitted to the liquid.

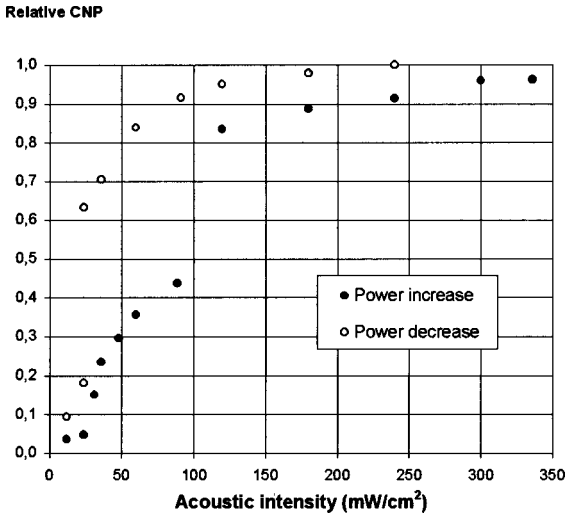


FIG. 7. Cavitation noise power as a function of the acoustic power transmitted to the liquid for increasing power sweep and for decreasing power sweep.

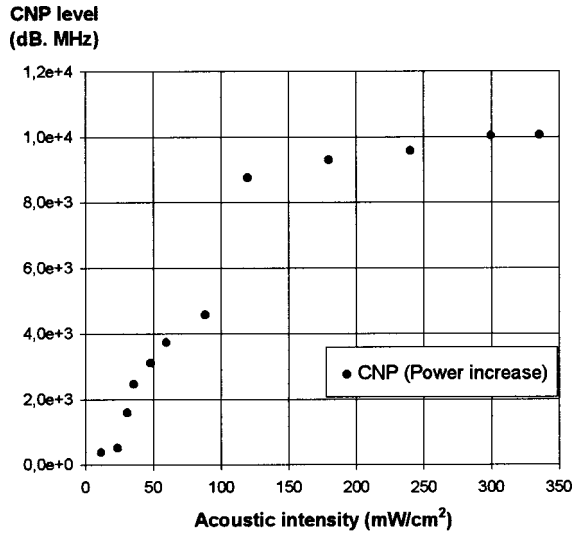


FIG. 5. Noise contribution to the cavitation noise power emission.

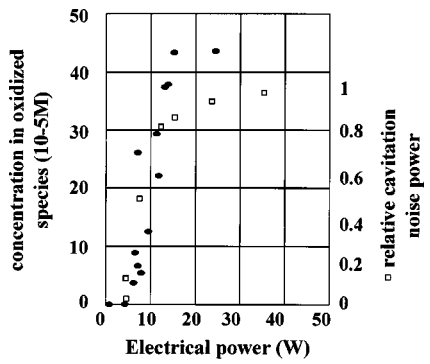


FIG. 8. Experimental correlation between acoustic cavitation noise power and sonochemical activity in an oxidization process.

noise is recorded with a DAPCO NP 10-3 hydrophone. The sonicated volume is 25 ml. The initial solution is 2.2×10^{-2} M(Fe^{2+}). Figure 8 represents the concentration of oxidized species after a 30-min sonication for different driving intensities. We have noticed a very good agreement between the cavitation noise measured and the oxidized species concentration variation as a function of the driving power. For our specific sample, we have found that the triggering of sonochemical reactivity is closely linked to the emergence of white noise in the cavitation noise spectrum.

VI. CONCLUSION

In this paper, we have first presented a simple model which allows one to link the $f/2$ oscillations to the well-sized bubble parametric resonance. Then, we have shown that in order to correctly characterize the complex cavitation phenomenon, the stable cavitation regime should be distinguished from the transient cavitation state. Some parameters such as $f/2$ and $2f$ peak levels are well-suited to characterize the first situation. We present a new tool adapted to identify the second one. Furthermore, due to the fluctuations induced by the chaotic behavior of the bubbles, the peak levels are more difficult to appreciate than the overall contribution of the cavitation noise. This powerful tool allows both cavitation characterization and sonochemical reactivity overview. Moreover, this new criterion combines real-time measurements and easy implementation. We have validated this probe in the case of an oxidization process and we have used it to show the hysteresis effect of cavitation. So, this method is suitable for cavitation characterization within industrial applications. However, further work will consist of applying this new tool to different industrial cavitation applications and to check if the previous results obtained near 1 MHz are valid in another frequency ranges.

[1] E. Graham and M. Hedges, ''Cavitational bio-effects at 1.85 MHz,'' Ultrasonics **18**, 224–228 (1980).
[2] G. Scherba and R. M. Weighel, ''Quantitative assessment of the germicidal efficacy of ultrasonic energy,'' Appl. Environ. Microbiol. **57**, 2079–2084 (1991).
[3] S. B. Barnett and M. C. Ziskin, ''Current status of research of biophysical effects of ultrasound,'' Ultrasound Med. Biol. **20**, 205–215 (1994).
[4] C. Sehgal, R. P. Steer, R. G. Sutherland, and R. E. Verrall, ''Sonoluminescence of argon saturated alcali metal salt solutions as a probe of acoustic cavitation,'' J. Chem. Phys. **70**, 2242–2248 (1979).
[5] K. S. Suslick, J. J. Gawienowski, P. F. Schubert, and H. H. Wang, ''Sonochemistry in nonaqueous liquids,'' Ultrasonics **22**, 33–36 (1984).
[6] K. S. Suslick, D. A. Hammerton, and R. E. Cline, ''The sonochemical hot spot,'' J. Am. Chem. Soc. **108**, 5641–5642 (1986).
[7] K. S. Suslick, S. J. Doktycz, and E. B. Flint, ''On the origin of sonoluminescence and sonochemistry,'' Ultrasonics **28**, 280–290 (1990).
[8] T. Lepoint, N. Voglet, L. Faille, and F. Mullie, ''Bubbles deformation and interface disruption as a source of sonochemical and sonoluminescent activity,'' in Bubble dynamics and interface phenomena, Proceedings of an IUTAM symposium, Birmingham, Sept. 1993, pp. 321–333.
[9] T. Lepoint, D. De Pauw, F. Lepoint-Mullie, M. Goldman, and A. Goldman, ''Sonoluminescence, an alternative electrohydrodynamic hypothesis,'' J. Acoust. Soc. Am. **101**, 2012–2030 (1997).
[10] E. A. Neppiras, ''Acoustic cavitation,'' Phys. Rep. **61**, 159–251 (1980).
[11] R. E. Apfel, *Methods of Experimental Physics*, edited by P. D. Edmonds (Academic, New York, 1981), Vol. 19, Chap. 7.
[12] R. E. Apfel, ''Acoustic cavitation prediction,'' J. Acoust. Soc. Am. **69**, 1624–1633 (1981).
[13] W. Hentschel and W. Lauterborn, ''New speed record in long series holographic cinematography,'' Appl. Opt. **23**(19), 3263–3265 (1984).
[14] W. Lauterborn and W. Hentschel, ''Cavitation bubble dynamics studies by high speed photography and holography. II,'' Ultrasonics **24**, 59–65 (1986).
[15] W. Lauterborn, I. Eick, and A. Philipp, ''Approaching bubble dynamics with lasers, holography and computers,'' in Bubble dynamics and interface phenomena, Proceedings of an IUTAM symposium, Birmingham, Sept. 1993, pp. 299–310.
[16] B. Gindroz, ''Comparison expérimentale de techniques de mesure de germes de cavitation,'' La houille blanche, n° **7–8**, p. 567 (1992).
[17] L. Van Wijngaarden, ''On the equations of motion for mixtures of liquid and gas bubbles,'' J. Fluid Mech. **33**, 465–474 (1968).
[18] R. E. Caflish, M. J. Miksis, G. S. Papanicolaou, and L. Ting, ''Effective equations for wave propagation in bubbly liquid,'' J. Fluid Mech. **153**, 259–273 (1985).
[19] M. J. Miksis and L. Ting, ''Effects of bubbly layers on wave propagation,'' J. Acoust. Soc. Am. **86**, 2349–2358 (1989).
[20] R. Y. Nishi, ''The scattering and absorption of sound waves by a gas bubble in a viscous liquid,'' Acustica **33**(2), 65–74 (1975).
[21] L. R. Gavrilov, ''Free gas content of a liquid and acoustical techniques for its measurement,'' Sov. Phys. Acoust. **15**(3), 285–295 (1970).
[22] H. Medwin, ''Counting bubbles acoustically: A review,'' Ultrasonics **15**, 7–13 (1977).
[23] C. Delebarre, I. Baquet, J. Frohly, E. Radziszewski, and P. Pernod, ''Dynamic characterization of high frequency ultrasonic cavitation,'' Proceedings Ultrasonics Symposium, pp. 723–726 (1993).
[24] F. W. Gibson, ''Measurement of the effect of air bubbles on the speed of sound in water,'' J. Acoust. Soc. Am. **48**, 1195–1197 (1970).
[25] R. A. Roy, S. I. Madanshetty, and R. E. Apfel, ''An acoustic backscattering technique for the detection of transient cavitation produced by microsecond pulses of ultrasound,'' J. Acoust. Soc. Am. **87**, 2451–2458 (1990).
[26] S. I. Madanshetty, R. A. Roy, and R. E. Apfel, ''Acoustic microcavitation: Its active and passive acoustic detection,'' J. Acoust. Soc. Am. **90**, 1515–1526 (1991).
[27] B. E. Noltingk and E. A. Neppiras, ''Cavitation produced by ultrasonics,'' Proc. Phys. Soc. **B63**, 674–685 (1950).
[28] J. G. Kirkwood and H. A. Bethe, Office of science research, Report 558, USA (1942).
[29] F. R. Gilmore, Hydrodynamics laboratory, Report 26-4, California Institute of Technology (1952).
[30] A. Prosperetti ''Nonlinear oscillation of gas bubbles in liquids: Steady states solutions,'' J. Acoust. Soc. Am. **56**, 878–885 (1974).
[31] A. Prosperetti, L. A. Crum, and K. W. Commander, ''Nonlinear bubble dynamics,'' J. Acoust. Soc. Am. **83**, 502–514 (1986).
[32] W. Lauterborn, ''Numerical investigation of nonlinear oscillations of gas bubbles in liquids,'' J. Acoust. Soc. Am. **59**, 283–293 (1976).
[33] H. G. Flynn, ''Physics of acoustic cavitation in liquids,'' in *Physical Acoustics*, Vol. 1B, edited by W. P. Mason (Academic, New York, 1964).
[34] C. C. Mei and X. Zhou, ''Parametric resonance of spherical bubble,'' J. Fluid Mech. **229**, 29–50 (1991).
[35] M. Minnaert, ''On musical air bubbles and the sound of running water,'' Philos. Mag. **16**, 235–248 (1933).
[36] L. Landau and E. Lifchitz, ''Physique théorique. Mécanique,'' Edition en langues étrangères, Moscou, pp. 114–116 (1960).
[37] W. Lauterborn and U. Parlitz, ''Methods of chaos physics and their application to acoustics,'' J. Acoust. Soc. Am. **84**, 1975–1993 (1988).
[38] V. I. Llyichev, L. V. Koretz, and N. P. Melnikov, ''Spectral characteristics of acoustic cavitation,'' Ultrasonics **27**, 357–361 (1989).
[39] V. I. Llyichev, L. V. Koretz, and N. P. Melnikov, ''Emission of sound by a stationary bubble under stochastic oscillations in liquid,'' Acoust. Phys. **40**(2), 232–236 (1994).
[40] I. Looten, J. Assaad, and J. Frohly, ''Electrical matching of power transducers for acoustic cavitation applications,'' J. Acoust. Soc. Am. **98**, 2922(A) (1995).
[41] R. Esche, Acustica **2**, AB208 (1952).
[42] S. M. Cordry, L. A. Crum, and R. A. Roy, ''Temperature-related effects in single-bubble sonoluminescence,'' J. Acoust. Soc. Am. **98**, 2921(A) (1995).
[43] D. F. Gaitan, L. A. Crum, C. C. Church, and R. A. Roy, ''Sonoluminescence and bubble dynamics for a single, stable, cavitation bubble,'' J.

Acoust. Soc. Am. **91**, 3166–3183 (1992).
[44] C. Delebarre, J. Frohly, I. Baquet, Z. Derouiche, and C. Bruneel, ''Acoustic signature of the cavitation noise,'' Proceedings Ultrasonics Symposium, pp. 719–722, Wien (1993).
[45] T. Lepoint, F. Mullie, P. Jaspard, I. Baquet, J. Frohly, and C. Delebarre, ''Experimental correlation between cavitation noise signatures and chemical reactivity in homogeneous sonochemistry,'' Proceedings Ultrasonics Symposium, pp. 727–730, Wien (1993).

A turbulence spectral model for sound propagation in the atmosphere that incorporates shear and buoyancy forcings

D. Keith Wilson
U.S. Army Research Laboratory, ATTN: AMSRL-CI-EP, 2800 Powder Mill Road, Adelphi, Maryland 20783

(Received 31 March 1999; accepted for publication 28 July 2000)

A three-dimensional model for turbulent velocity fluctuations in the atmospheric boundary layer is developed and used to calculate scattering of sound. The model, which is based on von Kármán's spectrum, incorporates separate contributions from shear- and buoyancy-forced turbulence. New equations are derived from the model that predict the strength and diffraction parameters for scattering of sound as a function of height from the ground and atmospheric conditions. The need is demonstrated for retaining two distinct scattering length scales, one associated with scattering strength and the other with diffraction. These length scales are height dependent and vary substantially with the relative proportions of shear and buoyancy forcing. The turbulence model predicts that for forward-scattered waves the phase variance is much larger than the log-amplitude variance, a behavior borne out by experimental data. A new method for synthesizing random fields, based on empirical orthogonal functions, is developed to accommodate the height dependence of the turbulence model. The method is applied to numerical calculations of scattering into an acoustic shadow zone, yielding good agreement with previous measurements. [S0001-4966(00)00811-0]

PACS numbers: 43.28.Fp, 43.28.Vd [LCS]

I. INTRODUCTION

Turbulent motions in the lowest 1 to 2 km of the earth's atmosphere (called the *boundary layer*) are produced primarily by a combination of wind shear and solar heating of the ground (Fig. 1). These motions span a broad spectrum of spatial scales from millimeters to kilometers.[1,2] As discussed by many recent authors,[3–10] calculations of sound propagation in this part of the atmosphere requires accurate representation of the turbulence spectrum. The complexity of the turbulence dynamics, however, makes development of appropriate turbulence models for propagation calculations challenging.

A further practical difficulty is that turbulence modeling for wave propagation calculations is driven by the need for fully three-dimensional (3-D) spatial models of the turbulence. Structure along the direction of propagation, as well as in the directions transverse, must be known. But the emphasis of atmospheric boundary-layer research has historically been on 1-D spectra recorded by point sensors.[11] Time series of the turbulent fluctuations are converted to approximate spatial series by assuming that the turbulence is ''frozen'' while it moves past the sensor at the mean wind speed. Unfortunately extrapolation of the 1-D measured spectra to three dimensions is far from straightforward. Some initial efforts in multidimensional modeling of atmospheric turbulence have been made by Kristensen *et al.*,[12] Mann,[13] Peltier *et al.*,[14] and Wilson.[15] Numerical turbulence simulation is provoking additional interest in multidimensional turbulence statistics within the atmospheric research community.[16]

Much of the early work on wave propagation in random media was based on 3-D Gaussian turbulence models.[17,18] Such models have the distinct advantage of analytical convenience and clear interpretation of the results. However, shortcomings in the Gaussian model became apparent with the advent and widespread acceptance of Kolmogorov's[19] hypotheses for turbulence statistics, according to which a broad spectral region, having energy proportional to the wave number raised to the $-\frac{5}{3}$ power, exists in high Reynolds number flows such as the atmosphere. This region of the spectrum, known as the *inertial subrange*, is not captured by Gaussian models. Consequently, Tatarskii,[20] Ishimaru,[21] and others formulated the theory of scattering by inertial-subrange turbulence. More recently, Ostashev[22] extended inertial-subrange scattering theory to accommodate turbulent velocity as well as scalar fluctuations.

For propagation of optical and other short wavelength electromagnetic radiation, the discrepancy between the meteorological state-of-the-art and the needs for propagation calculations has not been a pressing issue. This is because most scattering effects of interest result from interactions between the propagating waves and inertial-subrange turbulence. Within the inertial subrange, it is assumed that isotropy holds. This premise allows 3-D models to be developed systematically from 1-D measurements. Therefore knowledge of atmospheric boundary-layer statistics gained from point measurements can be used directly to determine the turbulence parameters in the inertial-subrange scattering theory. This was successfully accomplished nearly three decades ago.[23]

The needs for acoustic wave propagation calculations differ significantly from optics, however. Acoustic wavelengths of interest are about one million-fold longer. Recognizing that large eddies, belonging to the *energy-containing subrange* rather than to the inertial, would play a significant role in acoustic scattering, many past studies actually took a step backward in a historical sense by abandoning the Kolmogorov model in favor of the Gaussian one.[5,24,25] This

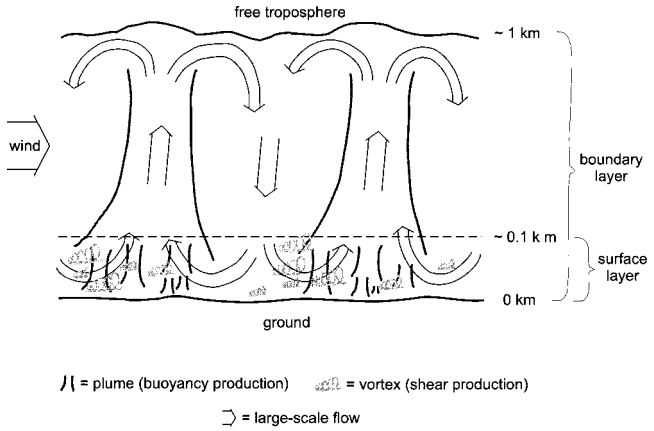


FIG. 1. Qualitative depiction of turbulence production and large-scale motions in the atmospheric boundary layer. Near the surface, turbulence is produced by shearing forces and plumes of rising warm air. The surface-layer plumes merge to create large circulations spanning the depth of the boundary layer.

course of action was justifiable, given that the Gaussian model is more appropriate to the energy subrange than Kolmogorov's.[10]

To reconcile the Kolmogorov and Gaussian approaches, recent investigators working in atmospheric acoustics[3,9,26] have suggested using a von Kármán model[27] for the turbulence spectrum. (The von Kármán spectrum has been used for several decades in electromagnetics.[28]) Parameters for the von Kármán model can be selected so that it agrees exactly with Kolmogorov's within the inertial subrange. And like the Gaussian model, von Kármán's has a realistic spectral roll-off at low wave number.

The main purpose of this article is to show how the von Kármán model can be systematically applied to scattering by atmospheric turbulence. The discussion is based on the framework of strength and diffraction parameters developed by Flatté *et al.*[29] (Sec. II). Determination of model parameters (Sec. III) is consistent with previous characterizations of turbulence spectra in the atmospheric boundary layer. The basic approach, which includes turbulence generated by both shear and buoyancy instabilities, is in fact nearly the same as that of Peltier *et al.*,[14] even though those authors considered only 2-D horizontal spectra. Furthermore, the differing effects of the two types of turbulence on a propagating sound wave are discussed in this article. New expressions for the velocity structure-function parameter and the scattering strength and diffraction parameters for scattering are derived (Sec. IV). For simplicity, the emphasis in this article is on propagation through turbulence in the absence of refractive gradients and ground reflections. However, the model can be applied to other refractive situations as well. In the final section (Sec. VI), it is applied to scattering into a ground-based refractive acoustic shadow zone.

Only turbulent wind velocity spectra are considered in this article, since the effect of velocity fluctuations is usually more important than that of temperature fluctuations, even when buoyancy forcing is the dominant mechanism for turbulence production.[30] Temperature fluctuations will generally be more important for propagation through a stably stratified flow, such as occurs near the ground on still nights.

II. STRENGTH AND DIFFRACTION PARAMETERS

The dimensionless strength and diffraction parameters proposed by Flatté *et al.*[29] conveniently characterize wave scattering in random media. The values for these parameters identify the propagation regime and, once the regime is known, can be used to calculate amplitude and phase variances of the received signal. The strength parameter Φ is simply the standard deviation of the acoustic phase fluctuations according to geometrical acoustics theory. Assuming that the length X of the propagation path is much longer than the integral length scale $\mathcal{L}$ parallel to the direction of propagation, the strength parameter is given by the equation[29]

$$\Phi^2 = 2\overline{\sigma^2} k_0^2 X \mathcal{L}, \tag{1}$$

where

$$\mathcal{L} = \frac{1}{\overline{\sigma^2}} \int_0^\infty \overline{R}(r\mathbf{e}_\parallel) dr, \tag{2}$$

where $\overline{\sigma^2}$ is the variance of the effective index-of-refraction fluctuations, $k_0 = \omega/c_0$ is the wave number, $\overline{R}(\mathbf{r})$ is the two-point correlation function ($\mathbf{r}$ being the separation between the measurement points), and $\mathbf{e}_\parallel$ is the unit vector in the direction of propagation.[31] For calculation of the effects of wind velocity fluctuations on a forward-propagating acoustic wave, the correlation function in Eq. (2) is that of the velocity fluctuation component v parallel to the direction of propagation, normalized by a reference sound speed c_0.[4] Overbars are used in this article to indicate statistics of v/c_0; absence of an overbar implies statistics of v alone.

The diffraction parameter Λ is proportional to the path-averaged ratio of the Fresnel zone radius to the correlation length L, the latter being defined by an expansion of the correlation function for small arguments perpendicular to the propagation direction:[29,32]

$$R(r\mathbf{e}_\perp) \simeq \sigma^2 \left(1 - \left| \frac{r}{L} \right|^{p-1} \right), \tag{3}$$

in which $\mathbf{e}_\perp$ is the unit vector perpendicular to the propagation and $p = \frac{5}{3}$ for turbulence. Based on this definition of L, the diffraction parameter is given by the equation[29,33]

$$\Lambda = \frac{X}{6L^2 k_0}. \tag{4}$$

The boundaries between the different scattering regimes, in terms of the strength and diffraction parameters, are illustrated by Fig. 2. Also indicated on the figure are the signal variances for the different regimes, where $\psi = \exp(\chi + i\phi)$ is the wave function phasor normalized by the mean amplitude, χ is the log-amplitude, and ϕ is the signal phase.

We see now that the main quantities of interest to be derived from the turbulence model (so that we can determine the strength and diffraction parameters) are the variance σ^2, the integral length scale $\mathcal{L}$, and the correlation length scale L of the turbulent velocity fluctuations. One might wonder

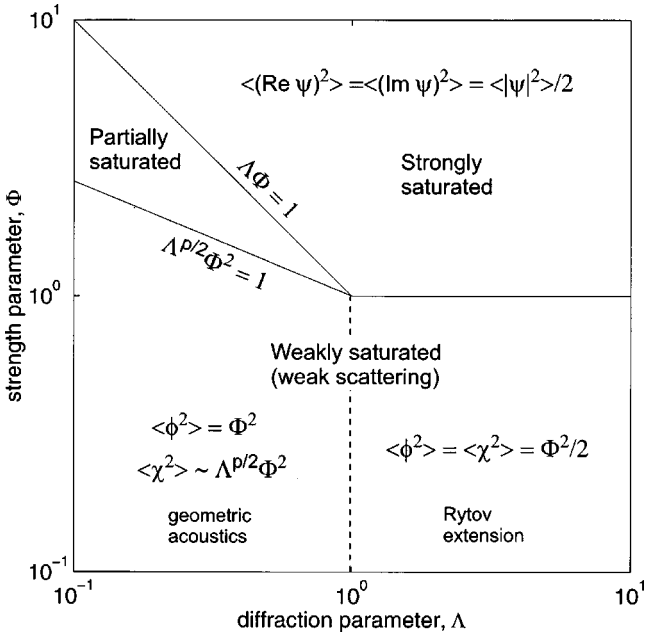


FIG. 2. Boundaries between the various scattering regimes in strength/diffraction space. The dashed line separates the weak scattering regime into a weak diffraction component, for which geometrical acoustics can be used to calculate the signal statistics, and a strong diffraction component, for which Rytov's method must be used to calculate the statistics. Statistics for the partial saturation regime are poorly understood. [This figure is adapted from Figs. 8.5 and 8.7 in Flatté *et al.* (Ref. 29).]

whether two separate length scales are really needed. Furthermore, how do these length scales relate to what is traditionally called the ''outer scale''? Indeed, there are well known connections between the correlation functions for parallel and perpendicular displacements appearing in Eqs. (2) and (3). As discussed by Batchelor,[34] the full autocorrelation function for an isotropic vector field is given by

$$R(\mathbf{r}) = \frac{r_\parallel^2}{r^2} R_\parallel(r) + \frac{r_\perp^2}{r^2} R_\perp(r), \tag{5}$$

where $\mathbf{r} = r_\parallel \mathbf{e}_\parallel + r_\perp \mathbf{e}_\perp$, $R_\parallel(r) \equiv R(r\mathbf{e}_\parallel)$, and $R_\perp(r) \equiv R(r\mathbf{e}_\perp)$. If the flow field is incompressible (as the turbulent velocity fluctuations are assumed to be),

$$R_\perp(r) = R_\parallel(r) + \frac{r}{2} \frac{dR_\parallel(r)}{dr}. \tag{6}$$

Therefore, for isotropic, incompressible turbulence, knowledge of $R_\parallel(r)$ leads to $R_\perp(r)$; both the integral and correlation length scales are known in principle once $R_\parallel(r)$ is specified. But this fact does not imply a simple, universal relationship between $\mathcal{L}$ and L. The following two sections of this article show that $\mathcal{L}$ and L each has its own distinctive dependence on the atmospheric forces producing the turbulence.

III. CHARACTERIZATION OF ATMOSPHERIC TURBULENCE

One important decision in modeling turbulence correlations and spectra is to decide which statistics of the turbulence will be used for the model parameters. In this article, I have chosen to follow Peltier *et al.*[14] by using two parameters for each component of the spectrum: the average variance of the velocity fluctuations, σ^2, and the dissipation rate of specific (per unit mass) turbulent kinetic energy (TKE), ϵ. These parameters have precise definitions, have been well studied in the atmospheric sciences literature, and can be determined with point sensors. Another reason for using the dissipation rate is that, according to Kolmogorov's hypotheses,[19] it determines the spectral density within the inertial subrange. The 1-D density for this part of the spectrum is[35]

$$F_\parallel(k) = \frac{\alpha}{2} \epsilon^{2/3} k^{-5/3}, \tag{7}$$

where α is a constant, whose value is approximately 0.52.[36] In this equation, the turbulent wave number k is implicitly parallel to the velocity fluctuations.

A. Scales for shear and buoyancy turbulence

Most descriptions of turbulence statistics, for the atmosphere as well as for other turbulent media, are based on dimensional scaling arguments. Several parameters, reasoned to be most significant for the flow dynamics, are selected and used to normalize the statistical quantities of interest. The appropriate parameters for atmospheric turbulence depend on the instability mechanism creating the turbulence. When surface-layer wind shear generates the turbulent energy, the appropriate length scale is z, the height from the surface, and the velocity scale is u_*, the *friction velocity*.[37] The friction velocity equals $\sqrt{\tau/\rho}$, where τ is the Reynolds stress exerted by the wind on the ground, and ρ is the ambient air density.

Atmospheric turbulence is also created by buoyancy force, which originates from solar heating of the ground and the consequent conduction of heat to a thin layer of overlying air. Pockets of warm air rising from the surface and compensating cool downdrafts from aloft combine to create large circulations encompassing the entire boundary layer, called *thermals* (Fig. 1). These circulations are constrained by the surface and an elevated layer having positive potential temperature gradient, called the *capping inversion*.[38] It has been well established that, in conditions of strong buoyancy production of turbulence, fluctuations in the near-ground horizontal velocity components originate primarily from these large circulations.[2,35] Hence the appropriate length scale for fluctuations in the near-ground horizontal wind components is the height of the capping inversion, z_i, which is typically several hundred to several thousand meters during the daytime. The appropriate velocity scale is the *mixed-layer velocity scale*, w_*,[1,39] given by $w_* = (gQ_s z_i/\theta_s)^{1/3}$, where Q_s is the surface temperature flux (units Km/s), θ_s is the surface temperature, and g is gravitational acceleration. (The heat flux H in W/m^2 can be calculated from Q_s using the relationship $H = \rho C_p Q_s$, where C_p is the specific heat at constant pressure.) In fair-weather daytime atmospheric boundary layers, buoyant production frequently dominates in the bulk of the boundary layer, although shear production dominates very near the ground.

Based on dimensional analysis, we therefore anticipate that the variance and dissipation rate for purely shear-driven turbulence are

$$\sigma_s^2 = a_s u_*^2 \quad \text{and} \quad \epsilon_s = b_s u_*^3/z, \tag{8}$$

where the subscript s means shear. For pure buoyancy turbulence (subscript b),

$$\sigma_b^2 = a_b w_*^2 \quad \text{and} \quad \epsilon_b = b_b w_*^3/z_i. \tag{9}$$

In the preceding equations, a_s, a_b, b_s, and b_b are dimensionless constants to be determined. These four constants constitute the ''unknowns'' for the von Kármán turbulence model presented later in this article. Values for them will be discussed in Secs. III B and C.

Many other turbulence scales commonly encountered in the meteorological literature can be written in terms of z, z_i, u_*, and w_*. Among the most significant of these is the Monin–Obukhov length, defined as $L_{\rm mo} \equiv -\theta_s u_*^3/\kappa g Q_s$, where $\kappa \simeq 0.40$ is von Kármán's constant. [The minus sign in the definition of $L_{\rm mo}$ is simply a convention, leading to $L_{\rm mo} < 0$ in unstable atmospheric conditions $(Q_s > 0)$.] Heights $z \ll -L_{\rm mo}$ are dominated by shear production of TKE, whereas heights $z \gg -L_{\rm mo}$ are dominated by buoyant production. From the definition of $L_{\rm mo}$, one can show

$$-\frac{z_i}{\kappa L_{\rm mo}} = \left(\frac{w_*}{u_*}\right)^3. \tag{10}$$

When the boundary-layer height z_i is much larger than $-L_{\rm mo}$, the structure of turbulence within the boundary layer as a whole is dominated by large, buoyantly produced eddies. Such boundary layers are termed *convective* in the atmospheric sciences literature. Conversely, when $z_i \ll -L_{\rm mo}$, the boundary layer is dominated by shear production. The latter situation is rather rare in the atmosphere.[1] In this article, I refer to the case $z_i \ll -L_{\rm mo}$ as a shear boundary layer, $z_i \sim -L_{\rm mo}$ as weakly convective, $z_i \sim -10 L_{\rm mo}$ as moderately convective, and $z_i \gtrsim -100 L_{\rm mo}$ as strongly convective.

B. Variances

Under conditions where shear production of turbulence dominates, the variances for the turbulent velocities are known to be anisotropic. Based on measurements of turbulence spectra made by Kaimal *et al.*,[35] Kristensen *et al.*[12] determined the variances $\sigma_1^2 = 4.8 u_*^2$, $\sigma_2^2 = 2.7 u_*^2$, and $\sigma_3^2 = 1.5 u_*^2$, where the subscript 1 indicates the along-wind direction, 2 the (horizontal) crosswind direction, and 3 the vertical. For an isotropic model, it is reasonable to set the variance equal to the average of all three velocity components. Hence, for shear turbulence, we have $a_s = 3.0$. (Application of complex anisotropic models to acoustic propagation through shear turbulence was discussed in an earlier paper.[15])

Under conditions where buoyant production of turbulence dominates the entire boundary layer, the velocity variances of the two horizontal components are equal and decrease somewhat with height, whereas the vertical velocity variance vanishes at ground level and increases with height.[40] This behavior is a consequence of blocking of the vertical motions by the ground and can be modeled accurately using potential flow theory.[41] Quite interestingly, the blocking appears to have no net effect on horizontal sound propagation.[42] For horizontal propagation, it is therefore reasonable to neglect the influence of the ground on the large eddies by setting $a_b = 0.35$, as suggested by Caughey and Palmer.[40]

Somewhat different values for a_s and a_b can be found from other sources in the literature, including Peltier *et al.*[14] The discrepancies result from scatter in the data and generally range between 10% and 30%.

C. Production and dissipation rates

A direct consequence of the logarithmic mean velocity profile in a turbulent shear layer is that the average rate of TKE production is $u_*^3/\kappa z$.[1,37] It is reasonable to assume for shear turbulence that TKE production at a given height is balanced by dissipation, implying $\epsilon = u_*^3/\kappa z$. Therefore $b_s = 1/\kappa$ in Eq. (8).

The rate of TKE production by buoyancy forces is $(g/\theta)\langle w'\theta'\rangle$, where w' and θ' are the fluctuation in vertical velocity and potential temperature.[1,38] The angle brackets indicate ensemble averaging. Since the vertical temperature flux $\langle w'\theta'\rangle$ decreases nearly linearly with height in a convective boundary layer, and Q_s is by definition the value of $\langle w'\theta'\rangle$ at the surface, the average dissipation rate of turbulence in the boundary layer must be proportional to w_*^3/z_i. Data analyzed by Caughey and Palmer[40] suggest that $\epsilon \simeq 0.8 w_*^3/z_i$ throughout most of a convective boundary layer, although ϵ increases somewhat near the ground, probably due to surface interactions. The value $b_b = 0.8$ is used in this article. In their model, Peltier *et al.*[14] set $b_b = 1$; existing data exhibit too much scatter to distinguish reliably between the two values.

D. Combining shear and buoyancy effects

When both shear and buoyancy production play significant roles in driving atmospheric turbulence (a common situation in the daytime boundary layer), a useful modeling assumption is that individual spectra representing the shear and buoyancy forcings can simply be added together. This idea apparently originated with Højstrup,[43] who had good success applying it to 1-D turbulence spectra. Peltier *et al.*[14] subsequently applied Højstrup's hypothesis to 2-D TKE spectra. Here it is assumed (as in Wilson and Thomson[5]) that the hypothesis is fully general, being valid for all 3-D velocity spectra and correlations.

By setting the argument of the correlation function to zero, it follows that the variances are also additive: $\sigma^2 = \sigma_s^2 + \sigma_b^2$. From Eqs. (8)–(10), one therefore finds

$$\begin{aligned}\sigma^2 &= a_s u_*^2 + a_b w_*^2 \\ &= u_*^2\left[3.0 + 0.35\left(-\frac{z_i}{\kappa L_{\rm mo}}\right)^{2/3}\right] \\ &= w_*^2\left[3.0\left(-\frac{z_i}{\kappa L_{\rm mo}}\right)^{-2/3} + 0.35\right].\end{aligned} \tag{11}$$

By integrating the correlation function with respect to the propagation direction, it can be shown that the squares of the strength parameters are also additive:

$$\Phi^2=\Phi_s^2+\Phi_b^2=\frac{2k_0^2X}{c_0^2}(\sigma_s^2\mathcal{L}_s+\sigma_b^2\mathcal{L}_b). \tag{12}$$

Integral length scales are variance-weighted sums:

$$\sigma^2\mathcal{L}=\sigma_s^2\mathcal{L}_s+\sigma_b^2\mathcal{L}_b\,. \tag{13}$$

As can be shown from Eq. (3), the correlation lengths combine in reciprocal fashion:

$$\frac{\sigma^2}{L^{p-1}}=\frac{\sigma_s^2}{L_s^{p-1}}+\frac{\sigma_b^2}{L_b^{p-1}}. \tag{14}$$

Limits on the applicability of results for the combined shear/buoyancy model, such as Eqs. (11)–(14), should be kept in mind. Equation (8) is restricted to the atmospheric surface layer, defined as $z<0.1z_i$.[38] Equation (9), on the other hand, is valid nearly to the top of the boundary layer (about $0.9z_i$). Therefore the combined model is generally valid below $0.1z_i$, although the region of validity should extend to $0.9z_i$ for moderately and strongly convective boundary layers, since the buoyancy contribution dominates shear in such circumstances.

IV. ISOTROPIC TURBULENCE MODELS

Strictly speaking, turbulence is never isotropic: forces generating the turbulence act directionally, and eddy structure in the production region of the spectrum (the energy-containing subrange) must reflect the generating mechanism. When the energy-containing subrange eddies break down into smaller ones, memory of the production mechanism can be lost, resulting in isotropic statistics for inertial-subrange scales (*local isotropy*) as hypothesized by Kolmogorov.[19] Still, since energy-containing subrange statistics are nonuniversal and very difficult to model, the assumption of isotropy across the entire turbulence spectrum is a practical and frequent starting point in statistical models. Furthermore, the assumption of isotropy often enables the analytical derivation of scattering statistics; without such analytical results there would be little physical insight into the scattering processes.

In this section, isotropic models for the turbulence spectrum are developed on the basis of the turbulence scales described in Sec. III. The inertial subrange model is discussed first, which serves to determine the correlation length scale needed for diffraction parameter calculations. A von Kármán model, incorporating both the energy-containing and inertial subranges, is described next. This model provides reasonable estimates of the integral length scale and, hence, the strength parameter. The section closes with a brief description of the Gaussian model, because this model has often been used in the past, and it will be compared to the von Kármán model in Sec. VI.

A. Inertial-subrange model

Kolmogorov's original hypotheses[19] imply that the structure function, defined as $D_\parallel(r)=\langle[v(r\mathbf{e}_\parallel)-v(\mathbf{0})]^2\rangle$, is proportional to $r^{2/3}$. Specifically,

$$D_\parallel(r)=C_V^2r^{2/3}, \tag{15}$$

where C_V^2 is the structure-function parameter for the velocity fluctuations. This equation is valid in the inertial subrange ($r/L\ll1$ or $kL\gg1$). Let us now consider relationships between C_V^2, L, and ϵ. From Eqs. (3) and (6), one can show that the correlation function for parallel displacements has the small-argument expansion

$$R_\parallel(r)\simeq\sigma^2\left(1-\frac{2}{p+1}\left|\frac{r}{L}\right|^{p-1}\right). \tag{16}$$

Since the definition of $D_\parallel(r)$ implies that $D_\parallel(r)=2[R_\parallel(0)-R_\parallel(r)]$, and $p=\frac{5}{3}$ for turbulence, we must have[44]

$$L=\left(\frac{3\sigma^2}{2C_V^2}\right)^{3/2}. \tag{17}$$

Turning now to the dissipation rate, it can be shown using the Fourier transform relationship between the correlation and spectral functions that

$$D_\parallel(r)=4\int_0^\infty(1-\cos kr)F(k)dk. \tag{18}$$

Note that wave numbers such that $kr\ll1$ do not contribute to the integral, because of the filtering action of the $(1-\cos kr)$ factor. Hence for small r we may use Eq. (7) for the spectrum. The integration can then be performed, resulting in

$$D_\parallel(r)=\tfrac{3}{2}\Gamma(\tfrac{1}{3})\alpha\epsilon^{2/3}r^{2/3}. \tag{19}$$

Comparing Eq. (19) to (15), we find

$$C_V^2=\tfrac{3}{2}\Gamma(\tfrac{1}{3})\alpha\epsilon^{2/3}\simeq2.1\epsilon^{2/3}. \tag{20}$$

This relationship is analogous to one derived for the temperature structure-function parameter by Wyngaard *et al.*[23]

Assuming the validity of Højstrup's hypothesis, one can find the overall C_V^2 by adding the contributions from shear and buoyancy turbulence. The result is

$$C_V^2\simeq\frac{3}{2}\Gamma\left(\frac{1}{3}\right)\alpha\left(\frac{b_s^{2/3}u_*^2}{z^{2/3}}+\frac{b_b^{2/3}w_*^2}{z_i^{2/3}}\right)\simeq3.9\frac{u_*^2}{z^{2/3}}+1.8\frac{w_*^2}{z_i^{2/3}}. \tag{21}$$

This result appears to be a new one, although Brown and Clifford[45] previously suggested using a relationship of the general form $C_V^2=A+B(z/z_i)^{-2/3}$, where A and B are constants. With regard to the correlation length, we have from Eqs. (17) and (20) with $\alpha=0.52$

$$L=\left[\alpha\Gamma\left(\frac{1}{3}\right)\right]^{-3/2}\frac{\sigma^3}{\epsilon}=0.61\frac{\sigma^3}{\epsilon}. \tag{22}$$

Combining the correlation lengths for shear and buoyancy turbulence in accordance with Eq. (14), the overall value L is found to be

$$L=\left[\alpha\Gamma\left(\frac{1}{3}\right)\right]^{-3/2}\left[\frac{a_su_*^2+a_bw_*^2}{b_s^{2/3}u_*^2/z^{2/3}+b_b^{2/3}w_*^2/z_i^{2/3}}\right]^{3/2}. \tag{23}$$

By factoring this expression in various ways and replacing the ratio u_*^2/w_*^2 with its equivalent in terms of the Monin–Obukhov length L_{mo} [Eq. (10)], L/z and L/z_i can be written entirely as functions of two of the following three dimensionless ratios: z/z_i, z/L_{mo}, and z_i/L_{mo}. For example, sub-

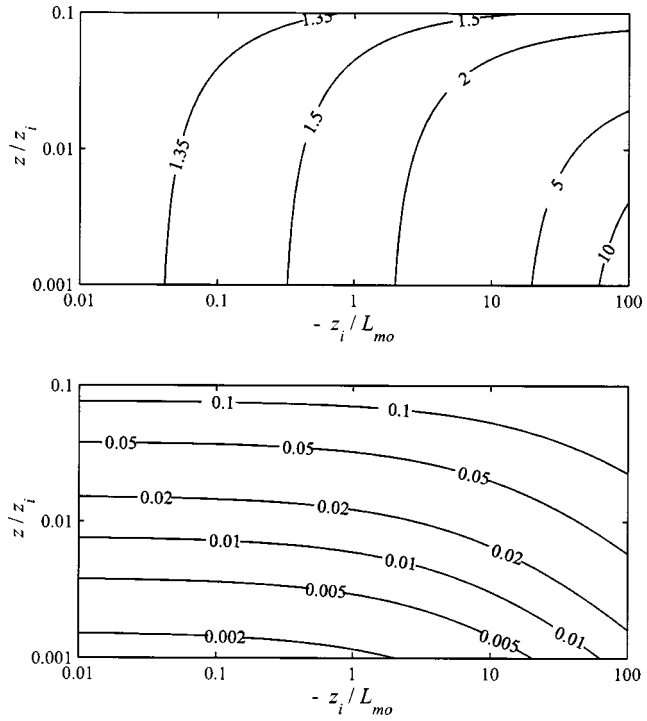


FIG. 3. Correlation length scale L as a function of $-z/L_{\mathrm{mo}}$ and z/z_i, where z is height from the ground, z_i is the boundary-layer inversion height, and L_{mo} is the Monin–Obukhov length. The upper plot shows L/z and the lower plot L/z_i.

stituting the numerical values from Secs. III B and C, we have

$$\frac{L}{z}=1.3\left[\frac{1+0.12(-z_i/\kappa L_{\mathrm{mo}})^{2/3}}{1+0.47(z/z_i)^{2/3}(-z_i/\kappa L_{\mathrm{mo}})^{2/3}}\right]^{3/2}, \tag{24}$$

or, alternatively,

$$\frac{L}{z_i}=0.16\left[\frac{1+8.5(-z_i/\kappa L_{\mathrm{mo}})^{-2/3}}{1+2.1(z/z_i)^{-2/3}(-z_i/\kappa L_{\mathrm{mo}})^{-2/3}}\right]^{3/2}. \tag{25}$$

When the boundary layer is dominated by shear production ($-z_i/L_{\mathrm{mo}}\ll 1$), the correlation length is simply $L_s=1.3z$. At the other extreme, when the boundary layer is dominated by buoyancy production ($-z_i/L_{\mathrm{mo}}\gg 1$) and the propagation height is above the shear layer ($-z/L_{\mathrm{mo}}\gg 1$), we have $L=L_b=0.16z_i$. The full behavior of L/z and L/z_i as functions of z/z_i and $-z_i/L_{\mathrm{mo}}$ is plotted in Fig. 3. The figures demonstrate that for propagation near the ground the correlation length, and therefore also the diffraction parameter, is determined primarily by shear turbulence unless conditions are strongly convective.

The strength parameter should not, strictly speaking, be calculated from the inertial-subrange correlation function. Equation (16) is valid only for small r/L; the integral length scale is nonconvergent if we substitute Eq. (16) directly into Eq. (2). In the turbulence literature (e.g., Ref. 37), the length scale $l=\sigma^3/\epsilon$ is often used to represent the large eddies and assumed to be close in value to the integral length scale. From Eq. (22), we therefore have $\mathcal{L}\simeq l\simeq 1.6L$. Flatté *et al.*[29] devise a different estimate for $\mathcal{L}$ in terms of L. However, such estimates cannot be generally valid. Determination of the actual relationship between $\mathcal{L}$ and L *requires information on the energy-containing subrange structure of turbulence.*

B. Von Kármán model

As discussed in the Introduction, the von Kármán model provides a simple, qualitative extension beyond the inertial subrange to the energy-containing subrange.[27,46] The model is actually presented in many different (and not always equivalent) forms in the literature. Here I define it based on the following equation for the TKE spectrum:[47]

$$E(k)=\frac{55}{9B(\frac{1}{2},\frac{1}{3})}\frac{\sigma^2k^4l^5}{(1+k^2l^2)^{17/6}}, \tag{26}$$

where σ^2 is the variance parameter (the variance of a single velocity component), l is a length scale parameter, and $B(x,y)=\Gamma(x)\Gamma(y)/\Gamma(x+y)$. To represent the dissipation subrange, some authors (particularly those concerned with optical propagation) use a more complicated equation with a spectral roll-off at high wave number. However, since the dissipation subrange has little effect on audible-frequency sound waves, it is neglected for this article. This simplification allows analytical derivation of equations for the correlation and structure functions. The correlation function for separations parallel to the direction of the velocity fluctuations determined using Eq. (26) is[47]

$$R_{\parallel}(r)=\frac{2\sigma^2}{\Gamma(\frac{1}{3})}\left(\frac{r}{2l}\right)^{1/3}K_{1/3}\left(\frac{r}{l}\right). \tag{27}$$

The 1-D longitudinal spectrum (the Fourier transform of the above correlation function) is

$$F_{\parallel}(k)=\frac{1}{B(\frac{1}{2},\frac{1}{3})}\frac{\sigma^2l}{(1+k^2l^2)^{5/6}}. \tag{28}$$

A reasonable approach for determining the parameters in the von Kármán model is to set σ^2 to the actual variance of the field and then choose l to match the Kolmogorov model in the inertial subrange. By taking the limit of Eq. (28) for large kl and setting the result equal to the right side of Eq. (7), one finds

$$l=\left[\frac{2}{\alpha B(\frac{1}{2},\frac{1}{3})}\right]^{3/2}\frac{\sigma^3}{\epsilon}\simeq 0.87\frac{\sigma^3}{\epsilon}. \tag{29}$$

Applying this equation to shear and buoyancy turbulence yields

$$l_s=0.87\,\frac{a_s^{3/2}}{b_s}z=1.8z \quad \text{and} \quad l_b=0.87\,\frac{a_b^{3/2}}{b_b}z_i=0.23z_i. \tag{30}$$

Since the von Kármán model reduces, by design, to the inertial-subrange model in the previous section when k is large, the correlation length scale L and the diffraction parameter Λ are the same for both models. Calculation of the strength parameter requires the integral length scale. A relationship between the integral length scale and l can be found by integrating Eq. (27), with the result

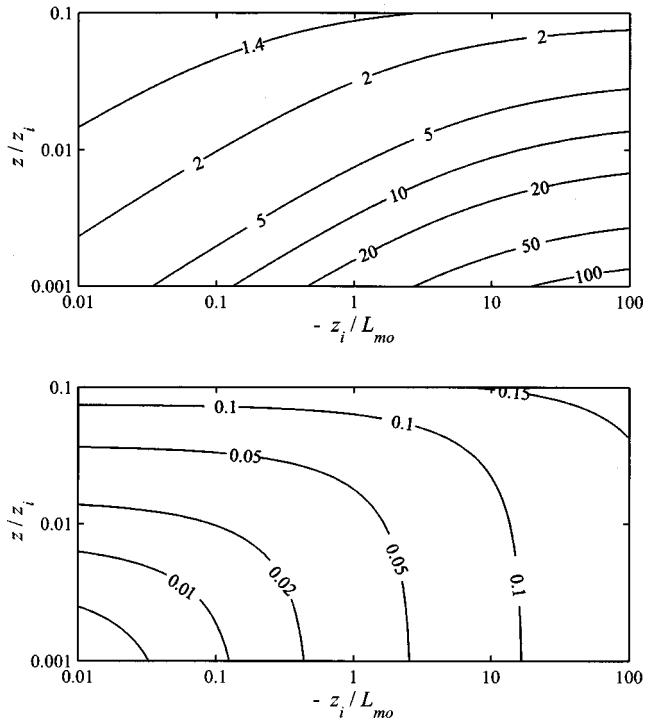


FIG. 4. Integral length scale $\mathcal{L}$ (parallel to the velocity fluctuations) as a function of $-z/L_{\mathrm{mo}}$ and z/z_i. The upper plot shows $\mathcal{L}/z$ and the lower plot $\mathcal{L}/z_i$.

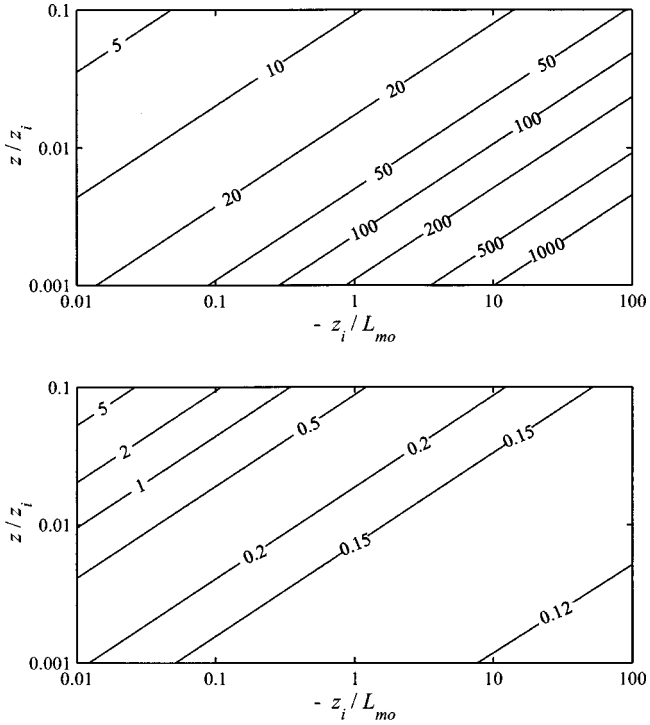


FIG. 5. Upper plot is the strength parameter normalized by the shear-turbulence scales, $\Phi c_0/(k_0 u_* X^{1/2} z^{1/2})$; lower plot is the strength parameter normalized by the buoyancy turbulence scales, $\Phi c_0/(k_0 w_* X^{1/2} z_i^{1/2})$.

$$\mathcal{L}=\frac{\pi}{B(\frac{1}{2},\frac{1}{3})}\,l\simeq 0.65\frac{\sigma^3}{\epsilon}. \tag{31}$$

According to Eq. (13), the general equation for $\mathcal{L}$ is found by weighting the length scales for shear and buoyancy turbulence by their respective variances and summing:

$$\mathcal{L}=0.65\frac{(a_s^{5/2}/b_s)u_*^2 z+(a_b^{5/2}/b_b)w_*^2 z_i}{a_s u_*^2+a_b w_*^2}. \tag{32}$$

Replacing the constants with their numerical values and rewriting in terms of the Monin–Obukhov length, we have

$$\frac{\mathcal{L}}{z}=1.3\frac{1+0.015(z_i/z)(-z_i/\kappa L_{\mathrm{mo}})^{2/3}}{1+0.12(-z_i/\kappa L_{\mathrm{mo}})^{2/3}} \tag{33}$$

and

$$\frac{\mathcal{L}}{z_i}=0.17\frac{1+67(z/z_i)(-z_i/\kappa L_{\mathrm{mo}})^{-2/3}}{1+8.5(-z_i/\kappa L_{\mathrm{mo}})^{-2/3}}. \tag{34}$$

Contour plots of $\mathcal{L}/z$ and $\mathcal{L}/z_i$ are shown in Fig. 4. Comparing Figs. 3 and 4, we observe no simple relationship between $\mathcal{L}$ and L over most of the range of $-z/L_{\mathrm{mo}}$ and z/z_i. On the contrary, there are very significant qualitative differences between $\mathcal{L}$ and L: near the ground, $\mathcal{L}$ usually is determined by the large buoyantly forced eddies, whereas L is determined by smaller shear-driven eddies. This holds true even for boundary layers with weak convection. We can view the distinctive behavior of $\mathcal{L}$ and L as resulting from the broad spectrum of spatial scales, spanning z to z_i, at which TKE is produced in the atmospheric boundary layer. The scale L falls near the boundary between the inertial and energy subranges, and hence is conceptually most like the ''outer'' scale used by researchers studying wave propagation. The scale $\mathcal{L}$ is weighted to the larger, more energetic eddies.

An expression for the normalized squared strength parameter follows from Eq. (12). Using the shear turbulence scales for normalization, one finds

$$\frac{\Phi^2 c_0^2}{k_0^2 u_*^2 X z}=7.9\left[1+0.015\left(\frac{-z_i}{\kappa L_{\mathrm{mo}}}\right)^{2/3}\frac{z_i}{z}\right]. \tag{35}$$

With the buoyancy scales,

$$\frac{\Phi^2 c_0^2}{k_0^2 w_*^2 X z_i}=0.12\left[1+67\left(\frac{-z_i}{\kappa L_{\mathrm{mo}}}\right)^{-2/3}\frac{z}{z_i}\right]. \tag{36}$$

Calculations of the strength parameter Φ from these equations are shown in Fig. 5. The strength parameter is dominated by buoyancy turbulence even when the boundary layer is only weakly convective. In Eq. (35), the initial term is the strength parameter for pure shear turbulence, $\Phi_s^2 \simeq 7.9 k_0^2 u_*^2 X z/c_0^2$. This value will be used in Sec. V to normalize several expressions for the statistics of scattered acoustic signals.

C. Gaussian model

The starting point for the Gaussian model is the longitudinal correlation function

$$R_{\parallel}(r)=\sigma^2\exp\left(-\frac{r^2}{l_{\mathrm{G}}^2}\right), \tag{37}$$

where l_{G} is the Gaussian length scale parameter. Fourier transformation yields the spectrum

$$F_{\parallel}(k)=\frac{\sigma^2 l_G}{2\sqrt{\pi}}\exp\left(-\frac{k^2 l_G^2}{4}\right). \tag{38}$$

Because it does not realistically describe the inertial subrange, the Gaussian model is best suited to the energy-containing subrange. However, it has also been used successfully as an ''applied filter'' when scattering occurs from only a limited part of the turbulence spectrum.[10]

Selection of l_G for scattering calculations is problematic. If the Gaussian correlation function is expanded for small argument, and the result set equal to Eq. (16) with $p=3$, then l_G is simply $\sqrt{2}L$. In this case, l_G is equivalent to a quantity known as the *Taylor microscale*.[10,37] In high Reynolds number turbulence, such as occurs in the atmospheric boundary layer, the Taylor microscale is orders of magnitude less than the integral length scale. Therefore, choosing l_G on the basis of the small-argument expansion leads to a very poor model for the inertial and energy subranges of the turbulence spectrum.[10]

A better procedure in many applications is to choose l_G on the basis of the integral length scale of the turbulence. The integral length scale for the isotropic Gaussian model is $l_G\sqrt{\pi}/2$. If the values of $\mathcal{L}$ in the von Kármán and Gaussian models are equated, one has the relationship

$$l_G=\frac{2\Gamma(\frac{5}{6})}{\Gamma(\frac{1}{3})}l\simeq 0.73\frac{\sigma^3}{\epsilon}. \tag{39}$$

The first part of this equation agrees with Eq. (6.51) in Ostashev.[22] This second part implies $l_G=1.5z$ for shear turbulence and $l_G=0.19z_i$ for buoyancy turbulence. These values differ somewhat from those suggested previously by Wilson and Thomson,[5] which were $1.0z$ and $0.40z_i$, respectively. For near-ground propagation, the value $l_G=1.5z$ agrees reasonably with the measurements of Daigle *et al.*[24] and Johnson *et al.*,[25] which were in the range of roughly 1 to 10 m.

V. CALCULATIONS OF SIGNAL VARIANCES AND COHERENCE FOR THE VON KÁRMÁN MODEL

A. Log-amplitude and phase variances

Ostashev[22] has analyzed the variances in log amplitude $\langle\chi^2\rangle$ and phase $\langle\phi^2\rangle$ for waves propagating in a medium with velocity fluctuations described by the von Kármán spectrum. For spherically propagating waves, the following equation [Ostashev's (7.66) and (7.67), valid for weak scattering ($\Phi\ll 1$)] derived with Rytov's method and the Markov approximation serves as the starting point:[21,22]

$$\langle\chi^2,\phi^2\rangle=\frac{\pi^2 k_0^2 X}{2}\int_0^1\int_0^\infty\left[1\mp\cos\frac{Xk_\perp^2 t(1-t)}{k_0}\right]\times\Phi_{\rm eff}(0,k_\perp)k_\perp dk_\perp dt, \tag{40}$$

where X is the propagation distance (range), $k_0=2\pi f/c_0$ is the acoustic wave number, f is the frequency, and $\Phi_{\rm eff}(k_\parallel,k_\perp)$ is the 3-D spectrum of the effective squared index-of-refraction fluctuations.[22] (The spectrum $\Phi_{\rm eff}$ should not be confused with the strength parameter Φ.) The magnitude of the turbulence wave numbers parallel and perpendicular to the direction of propagation are $k_\parallel$ and $k_\perp$, respectively. Based on Eq. (7.43) in Ref. 22 and Eq. (3.4.12) in Ref. 34, $\Phi_{\rm eff}(k_\parallel,k_\perp)=k_\perp^2 E(k)/(\pi c_0^2 k^4)$. We then have for the von Kármán spectrum given by Eq. (26)

$$\Phi_{\rm eff}(0,k_\perp)=\frac{55}{9\pi B(\frac{1}{2},\frac{1}{3})}\frac{\overline{\sigma^2}k_\perp^2 l^5}{(1+k_\perp^2 l^2)^{17/6}}. \tag{41}$$

Substituting this spectrum into Eq. (40) and changing variables to $u=k_\perp^2 l^2$, we find

$$\langle\chi^2,\phi^2\rangle=\frac{\Phi^2}{2}[1\mp f(D_l)], \tag{42}$$

where $D_l=X/l^2k_0$ and

$$f(D_l)=\frac{55}{36}\int_0^1\int_0^\infty\cos[D_l ut(1-t)]\frac{u}{(1+u)^{17/6}}du dt. \tag{43}$$

The solution to the double integral is unknown in analytical form. However, its limiting behavior for small and large D_l can be readily deduced. Due to the rapid oscillations in the cosine function, $f(D_l)$ approaches 0 as $D_l\to\infty$. Hence $\langle\chi^2,\phi^2\rangle=\Phi^2/2$ for large diffraction, as indicated in Fig. 2. For small D_l, the cosine function is approximately 1, and one finds

$$1-f(D_l)\simeq\frac{11}{6}\cos\frac{5\pi}{12}\Gamma\left(\frac{1}{6}\right)B\left(\frac{11}{6},\frac{11}{6}\right)D_l^{5/6}\simeq 0.582-D_l^{5/6}. \tag{44}$$

Therefore, for weak scattering and diffraction (geometrical acoustics regime), $\langle\phi^2\rangle=\Phi^2$. The log-amplitude variance is much smaller than the phase variance, being given by $\langle\chi^2\rangle\simeq 0.291-D_l^{5/6}\Phi^2$. This result, together with Eqs. (17), (20), (22), and (29), leads to

$$\langle\chi^2\rangle\simeq 0.708\Lambda^{5/6}\Phi^2\simeq 0.228(C_V^2/c_0^2)k_0^{7/6}X^{11/6}. \tag{45}$$

The form involving the strength and diffraction parameters is strictly valid only when the integral length scale prediction from the von Kármán model, Eq. (31), is used to calculate Φ^2. The form involving the structure-function parameter is not so limited and reduces to Tatarskii's[20] result for spherical spreading. (Actually, one must multiply Tatarskii's equation by $\frac{11}{6}$ and replace C_N^2 by C_V^2/c_0^2, since his result is for scalar index-of-refraction fluctuations, whereas the present result is for velocity fluctuations.[22])

Assuming as before that the shear and buoyancy contributions to the spectrum combine additively, Eq. (40) implies that the contributions to the log-amplitude and phase variances are also additive. Hence, starting with Eq. (42), and using (9) and (10) to rewrite the buoyancy parameters in terms of the shear parameters, one finds

$$\frac{\langle\chi^2,\phi^2\rangle}{\Phi_s^2}=\frac{1}{2}[1\mp f(D_{l,s})]+\frac{a_b^{5/2}b_s z_i}{2a_s^{5/2}b_b z}\left(-\frac{z_i}{\kappa L_{\rm mo}}\right)^{2/3}\times\left[1\mp f\left(\frac{a_s^3 b_b^2 z^2}{a_b^3 b_s^2 z_i^2}D_{l,s}\right)\right]. \tag{46}$$

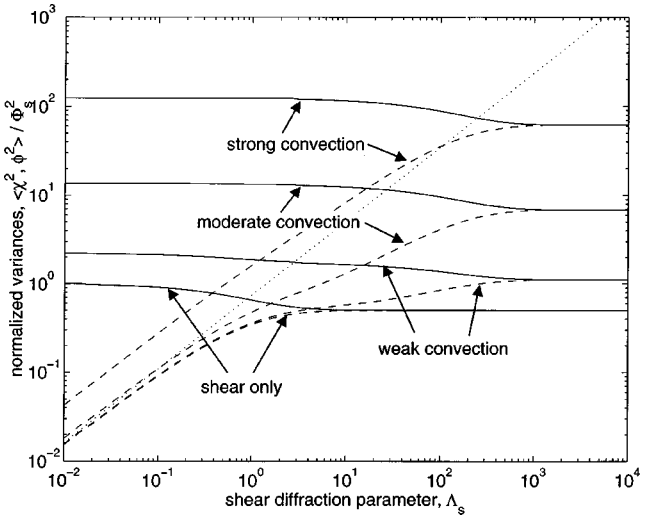


FIG. 6. Log-amplitude and phase variances for the von Kármán turbulence model for boundary layers having various degrees of convection (buoyancy forcing). The variances are normalized by the squared strength parameter for turbulence with pure shear forcing, Φ_s^2, and plotted as a function of diffraction parameter for pure shear forcing, Λ_s. The four solid and four dashed curves correspond to the phase and log-amplitude variances, respectively. Each set of four curves represents four different values of the ratio $-z_i/L_{\mathrm{mo}}$ (where z_i is the thickness of the boundary layer and L_{mo} is the Monin–Obukhov length): shear only, $-z_i/L_{\mathrm{mo}}=0$; weak convection, $-z_i/L_{\mathrm{mo}}=0.3$; moderate convection, $-z_i/L_{\mathrm{mo}}=10$; and strong convection, $-z_i/L_{\mathrm{mo}}=300$. The dotted line is the log-amplitude variance predicted by inertial-subrange theory in the presence of shear forcing only.

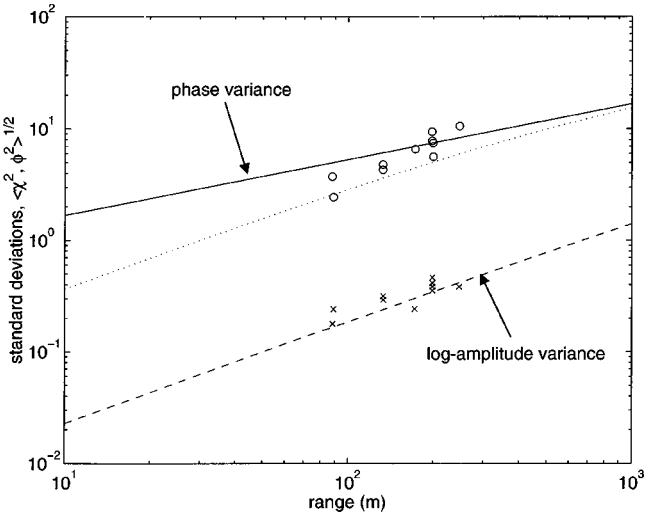


FIG. 7. Comparison of model predictions (solid and dashed curves) to standard deviations in phase (circles) and log-amplitude (×) measured by Wilson and Thomson. The dotted line is a numerical calculation of the phase deviation using geometrical acoustics, which (unlike the solid curve) is valid for short propagation distances relative to the size of the turbulent inhomogeneities.

This equation is plotted in Fig. 6 for $-z_i/L_{\mathrm{mo}}=0$ (shear only), 0.3 (weak convection), 10 (moderate convection), and 300 (strong convection). The ratio z/z_i was set to 0.01 in all the calculations. Also shown on the plot by the dotted line is the log-variance in the weak diffraction limit, Eq. (45), for the shear-only case. For small values of the shear diffraction parameter Λ_s (which equals $0.344D_{l,s}$), the log-amplitude variance is well described by the weak-diffraction, shear-only prediction. Only when convection is strong are significant deviations observed. The main role of increasing convection is to increase the phase variance relative to the log-amplitude variance. In very convective conditions, approximate equality of the phase and log-amplitude variances is obtained only for extremely large values of the shear diffraction parameter; these results indicate that the phase variance will always be much larger than the log-amplitude variance for acoustic frequencies less than several thousand Hz.

Previous experiments, such as Daigle *et al.*'s,[24] are in qualitative agreement with this prediction in that their measured phase variances are indeed larger than the log-amplitude variances. But direct comparisons between Eq. (46) and most previous experiments are problematic for three main reasons: (1) the experiments usually had short averaging times, which were insufficient to capture the slowly varying but strong phase fluctuations induced by the large, buoyantly forced eddies; (2) Eq. (46) strictly applies only to line-of-sight propagation, whereas ground reflections and refraction are significant in most datasets; and (3) concurrent meteorological data are insufficient to calculate parameters such as w_* and z_i. A small dataset that does not suffer from these limitations was collected by Wilson and Thomson[5] as part of their atmospheric acoustic tomography study. The signals in the experiment were swept sine waves instead of continuous tones, which allowed for separation of the direct and ground-reflected arrivals using a cross-correlation technique. The main drawback of these data for the present purpose is that the finite bandwidth of the signals (about 500 Hz) makes single-frequency calculations based on Eq. (46) tenuous.

The meteorological instrumentation operating during Wilson and Thomson's experiment included a cup anemometer at 10-m height, and bridged thermistors positioned at 2- and 10-m height. Using data from these instruments and a standard technique for fitting the mean wind and temperature profiles based on Monin–Obukhov similarity,[48] the values $u_*=0.35$ m/s and $L_{\mathrm{mo}}=-15/0.4=-38$ m were determined. Furthermore, $z_i=800$ m was measured using an optical ceilometer. Equation (10) then leads to $w_*=1.3$ m/s. The height of the speakers and microphones was $z=6$ m, and the nominal center frequency for the transmissions was $f=750$ Hz. Data from the experiment are compared to Eq. (46) in Fig. 7. Agreement is very good.

Despite the good results, applicability of Eq. (46) to Wilson and Thomson's tomography experiment is actually marginal for two reasons. First, the equation assumes that the propagation distance X is much larger than the integral length scale $\mathcal{L}$, which, based on Eq. (33), is 87 m for the experiment. To ascertain the validity of the approximation $X \gg \mathcal{L}$, Fig. 7 shows (by the dotted line) a numerical calculation of the standard deviation in phase that is based on the exact equation for geometrical acoustics [Flatté *et al.*'s[29] Eq. (6.2.4)], which is valid for any propagation distance. The slope of this curve indeed agrees with the phase data somewhat better than Eq. (46). Second, validity of the theory of weak scattering in this case is marginal. The calculated strength and diffraction parameters at 100 m are $\Phi=5.3$ and

$\Lambda = 0.022$, resulting in $\Lambda^{5/6}\Phi^2 = 1.2$. Hence the experiment actually lies near the transition between geometrical acoustics and partial saturation (Fig. 2), a region of the strength-diffraction space that is poorly understood.

B. Coherence

For plane waves, the coherence of the pressure field between two points $\mathbf{r}_1$ and $\mathbf{r}_2$, whose separation vector $\mathbf{r}=\mathbf{r}_1-\mathbf{r}_2$ is normal to the direction of propagation, is[21,22,47,49]

$$\Gamma^{pl}(\mathbf{r}) = \exp\{-\pi X k_0^2 \overline{d}_\parallel(\mathbf{r})\}, \tag{47}$$

where $d_\parallel(\mathbf{r})$, the *2-D structure function* of the velocity fluctuations parallel to the direction of propagation, is equal to

$$\overline{d}_\parallel(\mathbf{r}) = \int_0^\infty [1 - J_0(k_\perp r)]\Phi_{\text{eff}}(0,k_\perp) k_\perp \, dk_\perp \, . \tag{48}$$

The integration can be performed with $\Phi_{\text{eff}}(0,k_\perp)$ given by Eq. (41) and one finds[9,47,50]

$$b_\parallel(r) = \frac{2\sigma^2 l}{\sqrt{\pi}\Gamma(\frac{1}{3})}\left(\frac{r}{2l}\right)^{5/6}\left[K_{5/6}\left(\frac{r}{l}\right) - \frac{r}{2l}K_{1/6}\left(\frac{r}{l}\right)\right], \tag{49}$$

where

$$d_\parallel(\mathbf{r}) = 2[b_\parallel(\mathbf{0}) - b_\parallel(\mathbf{r})]. \tag{50}$$

Within the inertial subrange, $d_\parallel(\mathbf{r})$ depends only on the structure-function parameter, being simply[47]

$$d_\parallel(\mathbf{r}) = 0.850 C_V^2 r^{5/3}. \tag{51}$$

For the atmospheric model discussed earlier, Eq. (49) applies individually to the shear and buoyancy turbulence. To calculate the overall 2-D structure function, we add the 2-D correlations for shear and buoyancy [each represented by Eq. (49) with appropriate parameter values] and then apply the structure function definition, Eq. (50). Using the shear scales for normalization, the result is

$$\begin{aligned}\frac{d_\parallel(\overline{\mathbf{r}})}{zu_*^2} = {} & \frac{2a_s c_s}{B(\frac{1}{2},\frac{1}{3})}\left\{1 - \frac{2}{\Gamma(\frac{5}{6})}\left(\frac{\overline{r}}{2c_s z}\right)^{5/6}\left[K_{5/6}\left(\frac{\overline{r}}{c_s}\right)\right.\right. \\ & \left.\left. - \frac{\overline{r}}{2c_s}K_{1/6}\left(\frac{\overline{r}}{c_s}\right)\right]\right\} + \frac{2a_b c_b}{B(\frac{1}{2},\frac{1}{3})}\left(\frac{-z_i}{\kappa L_{\text{mo}}}\right)^{2/3} \\ & \times \frac{z_i}{z}\left\{1 - \frac{2}{\Gamma(\frac{5}{6})}\left(\frac{\overline{r}}{2c_b}\frac{z}{z_i}\right)^{5/6}\right. \\ & \left. \times \left[K_{5/6}\left(\frac{\overline{r}}{c_b}\frac{z}{z_i}\right) - \frac{\overline{r}}{2c_s z}K_{1/6}\left(\frac{\overline{r}}{c_s}\frac{z}{z_i}\right)\right]\right\},\end{aligned} \tag{52}$$

where $\overline{r} = r/z$ and $c_i = [2a_i/\alpha B(1/2,1/3)]^{3/2}/b_i$ ($c_s = 1.8$ and $c_b = 0.23$). Written in this form, $d_\parallel/zu_*^2$ depends only on $\overline{r}$, $-z/L_{\text{mo}}$, and $-z_i/L_{\text{mo}}$. (Similarly, an equation can be written for $d_\parallel/z_i w_*^2$ that depends only on r/z_i, $-z/L_{\text{mo}}$, and $-z_i/L_{\text{mo}}$.) Furthermore, with Eq. (47) the logarithm of the coherence can be written in the following normalized form analogous to Eq. (46):

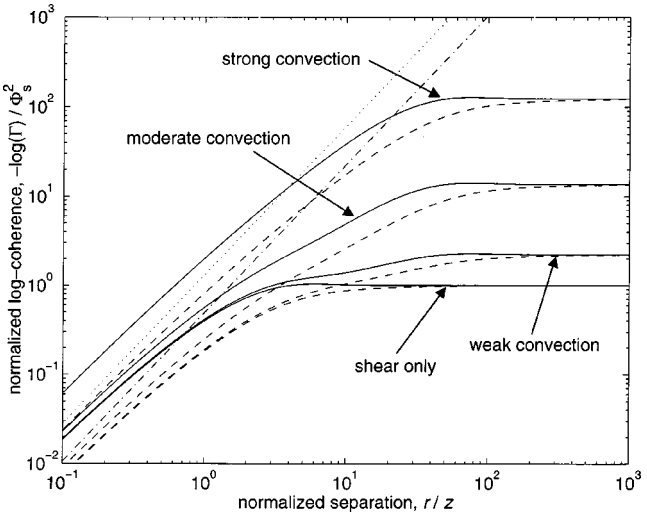


FIG. 8. Logarithm of the coherence function for the von Kármán turbulence model for boundary layers having various degrees of convection. The log coherence is normalized by the squared strength parameter for turbulence with pure shear forcing and plotted as a function of sensor separation r divided by height z. The four solid curves correspond to plane-wave propagation with the same four values of the ratio $-z_i/L_{\text{mo}}$ as in Fig. 6. The dashed curves are for spherical-wave propagation. The dotted line is the log coherence for plane waves predicted by Kolmogorov theory in the presence of shear forcing only. The dash-dotted line is the corresponding result for spherical waves.

$$-\frac{\log\Gamma^{pl}(\overline{r})}{\Phi_s^2} = \frac{B(\frac{1}{2},\frac{1}{3})}{2a_s c_s}\left[\frac{d_\parallel(\overline{\mathbf{r}})}{zu_*^2}\right]. \tag{53}$$

The log-coherence thus normalized by the strength parameter for shear turbulence is plotted in Fig. 8. The four solid curves represent the same four cases shown in Fig. 6, again with $z/z_i = 0.01$. Also shown in the figure are predictions from the Kolmogorov model for shear turbulence alone. An important observation is that for $r \ll z$, the log-coherence function always converges to the shear turbulence prediction unless convection is very strong. Shear turbulence alone is the main determinant of near-ground signal coherence for small sensor separations. For sensor separations larger than the height from the ground, however, the large convective eddies dominate the signal coherence. The strength and size of these eddies causes the coherence to be much less than for the pure shear-driven case.

For two measurement points both at a radial distance X from a point source, the coherence of a spherically propagating wave is[21]

$$\Gamma^{sp}(\mathbf{r}) = \exp\left\{-\pi X k_0^2 \int_0^1 \overline{d}_\parallel(\mathbf{r}t)\,dt\right\}. \tag{54}$$

The distinction between Eqs. (47) and (54) is that the structure function in (54) is averaged over the propagation path. The normalized log-coherence becomes

$$-\frac{\log\Gamma^{sp}(\overline{r})}{\Phi_s^2} = \frac{B(\frac{1}{2},\frac{1}{3})}{2a_s c_s}\left[\frac{\int_0^1 d_\parallel(\overline{rt})\,dt}{zu_*^2}\right]. \tag{55}$$

Calculations for spherically propagating waves are superimposed on the plane-wave results in Fig. 8. For small separations, the log coherence of the spherical wave is simply $\frac{3}{8}$

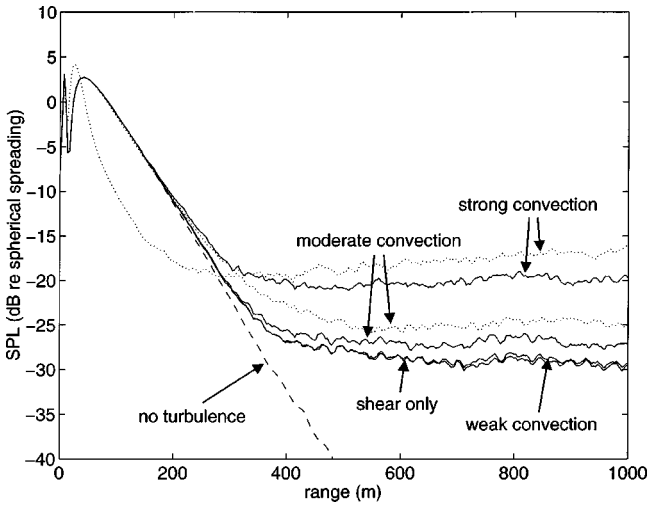


FIG. 9. Sound pressure levels relative to spherical spreading, for upwind propagation at $f=100$ Hz in boundary layers having various degrees of convection. The dashed line is the calculation without turbulent scattering. The solid lines represent turbulent scattering for varying degrees of atmospheric convection with a logarithmic sound speed profile. The dotted lines show calculations for more realistic sound speed profiles calculated using Monin–Obukhov similarity.

times the plane-wave value.[21] As the separation is increased, the spherical wave coherence gradually converges to the plane-wave result.

VI. CALCULATIONS OF SCATTERING INTO A SHADOW ZONE

Besides causing random fluctuations in received signals, scattering by turbulence is often the primary mechanism by which sound energy reaches a receiver in a refractive shadow zone.[51–53] In this section, the turbulence models developed earlier in this article are used to calculate mean sound pressure levels in the shadow. The mean levels are determined by running a parabolic equation (PE) code[54] for a large number of synthesized random turbulence fields, which are generated using the generalized random phase method discussed in the Appendix. The synthesized field consists of the horizontal wind velocity fluctuations in the nominal direction of acoustic propagation and is incorporated into the PE calculations using the phase screen method discussed by Gilbert *et al.*,[53] with phase screens placed at intervals of eight acoustic wavelengths.

A. Effect of atmospheric convection

The effect of varying degrees of atmospheric convection on sound levels in a refractive shadow zone is shown in Fig. 9. For this example, the acoustic frequency is $f=100$ Hz, and the source and receiver heights are 3.66 m. The normalized ground impedance is $Z/\rho_0 c_0 = 6.52 + i4.55$. For all the cases, the friction velocity is $u_* = 0.5$ m/s, and the inversion height is 1000 m. The mean refractive profile is the same for each of the cases, in order to isolate the effect of the turbulence structure. Specifically, the profile is logarithmic (strictly valid only for shear conditions), being $\langle c \rangle = c_0 + (u_*/\kappa)\ln(z/z_0)$ for $z > z_0$ and $\langle c \rangle = c_0$ for $z \leqslant z_0$, where $c_0 = 340$ m/s, and the roughness length z_0 is 0.01 m. The solid curves represent the same four values of the ratio $-z_i/L_{\mathrm{mo}}$ as in Sec. V, corresponding to $w_* = 0$, 0.454, 1.46, and 4.54 m/s for the shear only, weak convection, moderate convection, and strong convection cases, respectively. The sound level is nearly the same (about -30 dB relative to spherical spreading) for the shear-only and weak convection cases and only slightly larger for moderate convection. Strong convection increases the sound level deep in the shadow by about 10 dB.

One hundred random synthesized fields were averaged to obtain the results shown in Fig. 9. Example snapshots of the synthesized turbulence fields are shown in Figs. 10 and 11. The first of these is for shear turbulence only ($w_* = 0$ m/s), and the second is for strong convection ($w_* = 4.54$ m/s). When only shear is present, the ''structures'' are clearly smallest near the ground, gradually merging into larger ones at greater heights. Strong convection results in stronger wind speed fluctuations overall and larger structures near the ground.

In actuality, the logarithmic profile used in Fig. 9 is unrealistic for convective conditions. More realistic refractive profiles can be derived using Monin–Obukhov similarity, as discussed by L'Espérance *et al.*[48] The similarity technique requires a value for $T_* = -Q_s/u_*$, the surface-layer temperature scale. In the example here, T_* can be determined using the relationship $L_{\mathrm{mo}} = \theta_s u_*^2/\kappa g T_*$, resulting in $T_* = 0$, 0.0057, 0.19, and 5.7 K for the four values of z_i/L_{mo} considered. The two dotted lines in Fig. 9 show the sound levels for moderate and strong convection when the more realistic refractive profiles are used. (The corrections for the other two cases are negligible.) For moderate convection, the sound levels deep in the shadow are 2 to 3 dB higher than for the logarithmic profile. For strong convection, the sound levels near the source (before the shadow) are much lower, but gradually increase to about 4 dB higher at 1000 m from the source. The reason why the levels near the source are so low is that there is a very strong negative temperature gradient near the surface, which intensifies upward refraction of the sound.

The results shown in Fig. 9 can be understood on the basis of a simple Bragg scattering analysis, as suggested by Stinson *et al.*[55] The nearly forward scattering is analogous to a diffraction grating with periodic structures having a vertical spacing of D. Constructive interference of the scattered waves occurs when the acoustic wavelength is $2D\sin(\theta/2)$, where θ is the scattering angle. Assuming for illustrative purposes that the scattering occurs midway between the source and receiver at a height of 20 m, and that the horizontal separation between the source and receiver is 1000 m, one finds that the vertical wave number of the turbulence for resonant scattering is $k_{\mathrm{B}} = 0.12\ \mathrm{m}^{-1}$. In actuality, the scattering comes from a spectral band in the vicinity of k_{B}.[8] The vertical spectrum of the turbulence can be found by calculating $R_\perp(r)$ from Eqs. (6) and (27) and then taking the Fourier transform. The result is

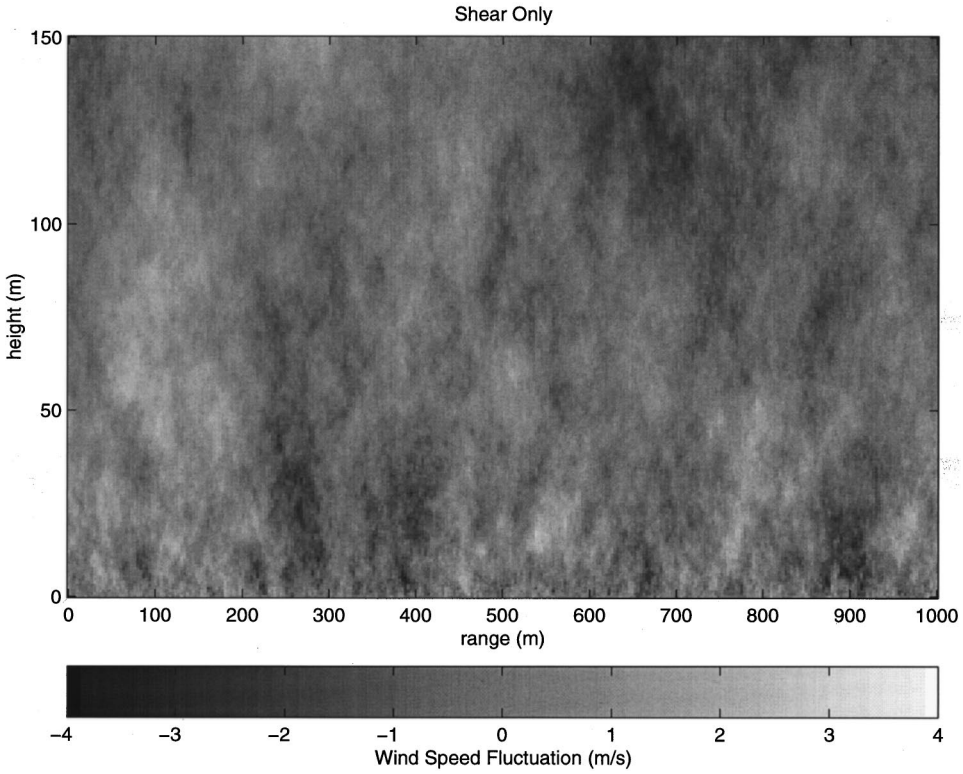


FIG. 10. Synthesized turbulence field with shear forcing only ($u_*=0.5$ m/s, $w_*=0$ m/s). The length scale is proportional to the height, and the variance is height independent.

$$F_\perp(k)=\frac{1}{B(1/2,1/3)}\,\frac{\sigma^2 l}{(1+k^2l^2)^{5/6}}\left[\frac{4}{3}-\frac{5}{6(1+k^2l^2)}\right]. \qquad (56)$$

To find the overall spectrum, we sum the $F_\perp(k)$'s for shear and buoyancy turbulence. The result is shown in Fig. 12, with a vertical dashed line marking $k_B=0.12\ \mathrm{m}^{-1}$. Note that k_B falls near the beginning of the inertial subrange. Hence, at least for this source/receiver geometry, using the inertial-subrange model for scattering into the shadow zone is justifiable for $f=100$ Hz and higher frequencies. Indeed, comparing Figs. 9 and 12, increases in the inertial-subrange spectral level are observed to cause proportionate increases in the sound level deep within the shadow. Frequencies below 100 Hz will interact to a greater extent with the energy-containing subrange turbulence, necessitating accurate modeling of that part of the spectrum.

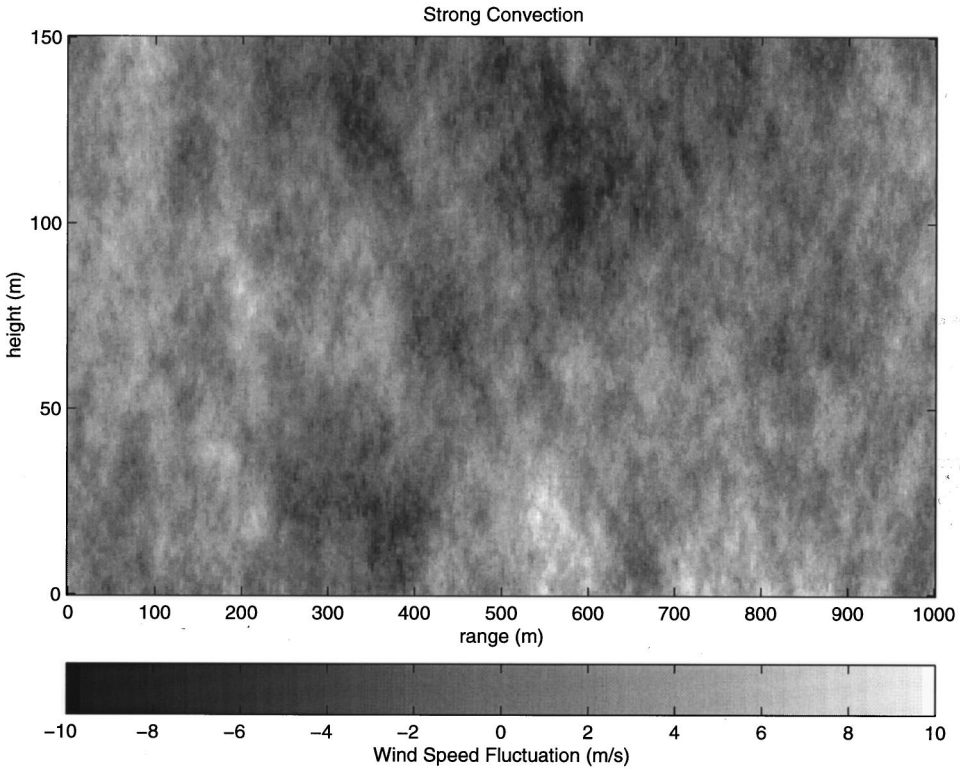


FIG. 11. Synthesized turbulence field with strong buoyancy forcing ($u_*=0.5$ m/s, $w_*=4.54$ m/s). The length scale and variance for the buoyantly forced component of the turbulence are both height independent.

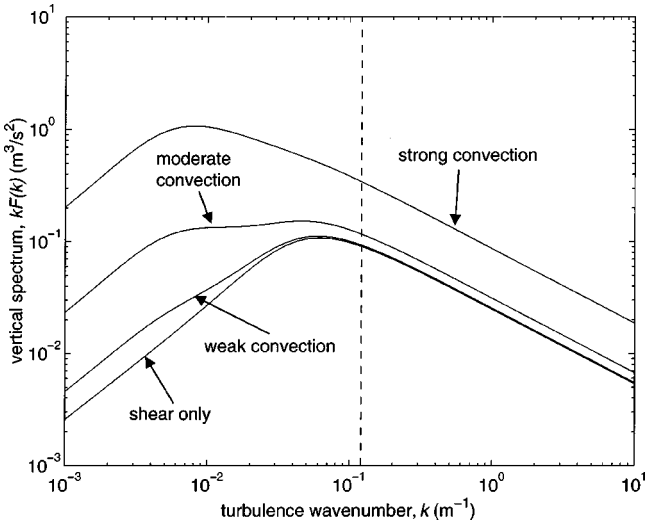


FIG. 12. Modeled vertical spectrum of atmospheric turbulence for varying degrees of convection. The dashed line indicates the Bragg scattering wave number for scattering from a point 20 m high and midway between the source and receiver, which are 1000 m apart.

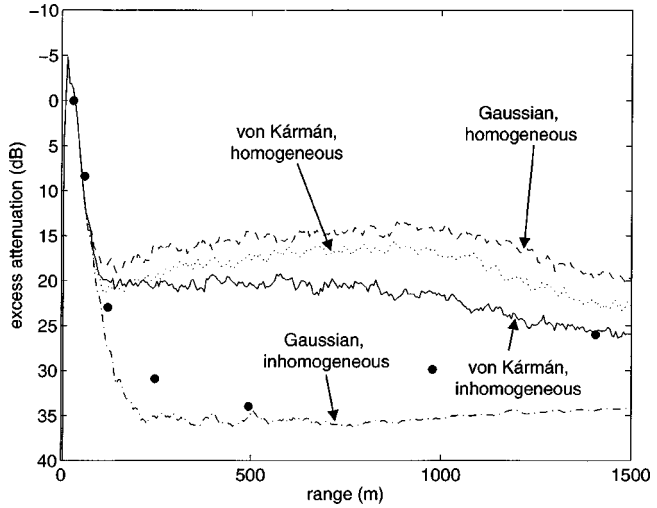


FIG. 13. Comparison of Wiener and Keast's excess attenuation measurements (circles) for propagation nearly upwind to predictions from four turbulence models.

B. Comparison to Wiener and Keast's data

Several decades ago, Wiener and Keast[51] collected an extensive set of measurements of outdoor acoustic attenuation. Their data for upwind propagation have since been used frequently to test theories and numerical calculations of sound levels in shadow regions.[3,53] Gilbert *et al.*[53] were able to obtain fair agreement with Wiener and Keast's data using a logarithmic sound-speed profile and a Gaussian turbulence model with a height-independent length scale. Juvé *et al.*[3] found that a von Kármán model with a height-independent (constant) length scale improved the agreement somewhat. But as discussed in Secs. IV C and VI A, both the logarithmic profile and height-independent length scale are unrealistic. To test the inhomogeneous von Kármán turbulence model proposed in this article, at a minimum it should be demonstrated that the new model also agrees satisfactorily with Wiener and Keast's data.

Rigorous comparisons to Wiener and Keast's data are made difficult by a lack of information on some key aspects of the meteorological conditions during their measurements. Although Wiener and Keast assembled good meteorological instrumentation by the standards of the day, they did not measure the boundary-layer inversion height and did not publish values for the surface heat flux. The important roles of these parameters in determining the mean profiles and turbulence structure were not widely appreciated among boundary-layer meteorologists during the 1950s. So, in making comparisons to their acoustic data, I have had to make some reasonable assumptions. First, since the results in the previous section of this article indicate that scattering into a shadow region from buoyantly forced turbulence is normally small except for strong convection, the contribution from the buoyancy turbulence is neglected. This assumption negates the need for information on the inversion height. Second, rough estimates of the surface heat flux are made in order to calculate the mean wind and temperature profiles. For their measurements taken almost directly upwind (the upper half of Fig. 2 in their paper[51]), Wiener and Keast state that skies were clear and there was a lapse temperature profile. A modest value for the surface heat flux during these measurements is therefore assumed, namely $H=200$ W/m^2 ($Q_s=0.17$ Km/s). (The heat flux was discussed in Sec. III A.) For their measurements taken nearly crosswind (the lower half of their Fig. 2), Wiener and Keast call the conditions neutral and overcast. Therefore for this case H is set to a very small value: 10 W/m^2 ($Q_s=0.0083$ Km/s).

Once the surface heat flux has been specified, a nonlinear equation for the friction velocity [Eq. (3) in L'Espérance *et al.*[48]] can be solved if the wind speed is known at some height. Wiener and Keast measured wind speeds of 15 and 17 MPH at 10-ft height (6.7 and 7.6 m/s at 3.05-m height) for their two cases. Based on this information, the values $u_*=0.50$ and 0.54 m/s are determined for the two cases. The corresponding surface-layer temperature scales now follow from the definition $T_*=-Q_s/u_*$, and the mean wind and temperature profiles can be calculated using Eqs. (3) and (4) in L'Espérance *et al.*[48]

The PE predictions from various turbulence models are compared to Wiener and Keast's data for the 300- to 600-Hz band in Figs. 13 and 14. The source height is 3.94 m and the receiver height 1.64 m. Plotted is the attenuation in excess of spherical spreading, with the reference (0 dB) at 100-ft (30.5-m) horizontal distance from the source. For the PE predictions, the precedent of Gilbert *et al.*[53] was followed by setting the frequency to 424 Hz. Four turbulence models are considered: the von Kármán and Gaussian, each in homogeneous (height-independent length scale) and inhomogeneous (height-dependent length scale) forms. For the inhomogeneous von Kármán and Gaussian models, parameters were selected as described in Secs. IV B and C. The homogeneous Gaussian model has $\overline{\sigma^2}=2\times10^{-6}$ and $l_G=1.1$ m, as originally suggested by Daigle[56] and used subsequently by Gilbert *et al.*[53] The variance for the homogeneous von Kármán model is also $\overline{\sigma^2}=2\times10^{-6}$ and the length scale follows from l_G using Eq. (39).

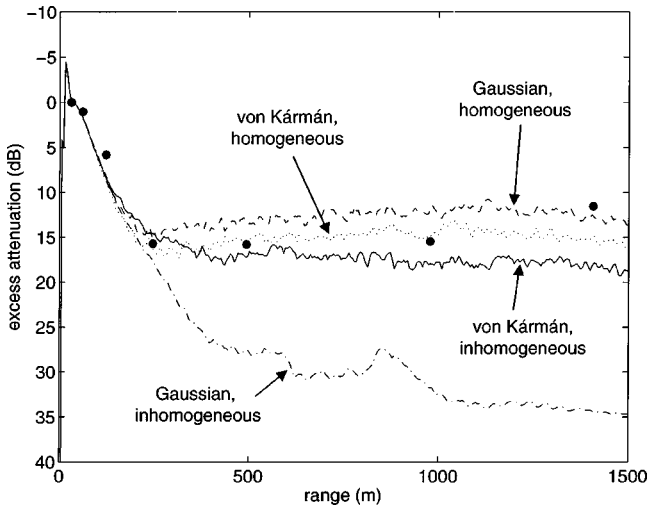


FIG. 14. Comparison of Wiener and Keast's excess attenuation measurements (circles) for propagation nearly crosswind to predictions from four turbulence models.

The measurements for nearly direct upwind propagation (Fig. 13) actually agree best with the inhomogeneous Gaussian model for distances less than 1 km, and then approach the predictions of the inhomogeneous von Kármán model (the one viewed here as most realistic) at 1.5 km. The homogeneous models both significantly underpredict the attenuation. For nearly crosswind propagation (Fig. 14), the homogeneous models and the inhomogeneous von Kármán model all agree reasonably with the data. The inhomogeneous Gaussian model overpredicts the attenuation by about 20 dB at 1 km.

It is difficult to draw any substantial conclusions from these comparisons. In both cases, the inhomogeneous von Kármán model predicts scattered signal levels about 5 to 10 dB less than the homogeneous Gaussian model that has been frequently used in the past. The inhomogeneous Gaussian model likely predicts unrealistically weak scattering. More detailed acoustical measurements, with supporting atmospheric data, are obviously needed to distinguish between the various turbulence models for scattering into a shadow region.

VII. CONCLUDING REMARKS

In the formulation developed by Flatté *et al.*,[29] two basic length scales characterize scattering in a random medium. One of these, the correlation length L, is associated with diffraction, and the other, the integral length scale $\mathcal{L}$, with geometric phase fluctuations (the ''strength'' of the index-of-refraction fluctuations). For scattering by turbulent velocity fluctuations, the scale L is determined by the total TKE dissipation rate, which in the near-ground atmosphere depends mostly on shear-generated eddies having size comparable to the height from the ground. In contrast, the scale $\mathcal{L}$ is weighted toward the largest, most energetic eddies, which in the atmosphere are usually produced by buoyancy forces and have sizes on the order of the overall depth of the atmospheric boundary layer. Hence L is much smaller than $\mathcal{L}$ when there is significant buoyancy forcing. Such conditions generally occur during the daytime unless there is a thick cloud layer and or very strong wind. The main manifestations of this disparity between the scattering length scales are that the phase variance for line-of-sight propagation is typically much larger than the log-amplitude variance, and that the signal coherence between a pair of microphones will typically decrease until they are more than 100 m apart. The necessity of defining two separate length scales was pointed out previously by Wilson and Thomson.[5] In that paper, the two length scales were referred to as the *parallel* length scale, $L_{\parallel}$, and the *perpendicular* length scale, $L_{\perp}$. These two scales are analogous to $\mathcal{L}$ and L, respectively, in this article.

This article has provided some new parametrizations for the length scales L and $\mathcal{L}$, the structure-function parameter C_V^2, and the variance $\overline{\sigma^2}$ of the velocity-induced fluctuations in the index of refraction. The parametrizations are based on Højstrup's hypothesis as applied to von Kármán's energy spectrum, although L can be determined on the basis of Kolmogorov theory alone. Second-order statistics of a scattered acoustic signal follow from knowledge of L, $\mathcal{L}$, and $\overline{\sigma^2}$. The parametrizations make use of the turbulence scales commonly employed by boundary-layer meteorologists, involving only the height of the propagation path, the boundary-layer inversion height, the friction velocity, and the mixed-layer velocity scale. The mixed-layer velocity scale can be calculated from the surface heat flux or the Monin–Obukhov length if the other turbulence scales are known. Methods for determining the friction velocity and Monin–Obukhov length based on the mean wind and temperature profiles near the ground are well established in boundary-layer meteorology. The inversion height can often be estimated with a ceilometer (an instrument for measuring the altitude of the cloud base) or radiosonde, or it can be estimated using numerical weather models.

The isotropic von Kármán turbulence model proposed in this article builds in a natural way on existing knowledge and methods for modeling atmospheric turbulence. It is practical for acoustic propagation calculations and captures the relevant turbulence structure in a large range of atmospheric conditions. Although the model is quite different from the homogeneous (constant length scale) Gaussian model now often used in atmospheric acoustics, its predictions are consistent with measurements of shadow-zone sound levels that have been previously explained using the homogeneous Gaussian model.

The von Kármán model as presented here does have some shortcomings and limitations that should be kept in mind. One aspect of realistic turbulence structure it does not capture is anisotropy. Anisotropy was treated in a previous paper,[15] where calculations suggested that stretching of eddies by shear can significantly affect acoustic propagation. One could, if desired, replace the shear-turbulence contribution to the isotropic model in this article with Mann's[13] anisotropic model. To do so would significantly increase complexity, however, making analytical solutions for quantities such as the 2-D structure function unobtainable.

Some additional modeling issues warranting consideration are the following: (1) The extent to which Højstrup's hypothesis is valid. Since turbulence is nonlinear, buoyancy

and shear forces must interact, and the simple addition of shear and buoyancy spectra must be a simplification. (2) Other mechanisms for generating large eddies besides buoyancy, such as the inflection-point instability.[57] (3) Nonstationarity, for example around sunrise and sunset. (4) The statistics of stable atmospheric boundary layers. Such boundary layers form on clear, still nights, and in these conditions the random fluctuations in the medium originate primarily from buoyancy (gravity) waves, rather than turbulence.

ACKNOWLEDGMENTS

This article is dedicated to John Wyngaard. Weaved through its fabric are many of the things he taught me about turbulence during my tenure as a research associate in the Pennsylvania State University Department of Meteorology. Discussions with J. Brasseur, K. Gilbert, J. Mann, D. Norris, V. Ostashev, and D. Thomson also played an essential role in formulating the ideas presented here.

APPENDIX: SYNTHESIS OF AN INHOMOGENEOUS RANDOM FIELD

During the past decade, acoustic propagation modeling has advanced to the point where full-wave solutions can be computed for propagation through a medium having 2-D and 3-D spatial variability in the index-of-refraction field.[53,58,59] Since a turbulence field cannot be directly measured to the fidelity needed for the wave propagation calculations, some method for synthesizing turbulence from a spectral model is desirable. Gilbert *et al.*[53] originally synthesized a turbulence field for acoustic calculations in a vertical 2-D plane by multiplying the square root of the spectrum at each wave number by a random phase, and then applying an inverse Fourier transform. Their method, which is similar to techniques used in other areas of research, has since been widely adopted in atmospheric acoustics. Henceforth I will call it the *random-phase method* (RPM). It is computationally inexpensive and produces synthetic turbulence having second-order statistics consistent with the assumed turbulence spectrum. However, the fields created by this method do not necessarily have third- and higher-order statistics that are realistic for turbulence. In particular, RPM cannot capture intermittent features known to be characteristic of turbulence.[60] But if one is interested only in calculating means and second-order statistics of the acoustic field, a method for producing synthetic turbulence that is realistic only to second-order may be satisfactory.

Another limitation of RPM as previously implemented is that the turbulence model was homogeneous. As discussed in Sec. III, the integral length scale for shear turbulence is proportional to height, making the statistics inhomogeneous in the vertical direction. Therefore we would like to extend RPM to synthesize inhomogeneous turbulence structure. Such a method, called the *generalized* random-phase method (GRPM), is introduced in this appendix. The underlying idea of GRPM is to multiply the empirical orthogonal functions (EOFs) by random phases and then sum them to form the synthetic turbulence field. The reason this method reduces to standard RPM in the case of homogeneous turbulence is that the EOFs of a homogeneous correlation function are harmonic functions.[61,62] The EOFs are defined as the eigenfunctions of the correlation function:

$$\int R(\mathbf{x},\mathbf{x}')\,\psi(\mathbf{x}')\,d\mathbf{x}' = \lambda\,\psi(\mathbf{x}), \tag{A1}$$

where $\mathbf{x}=(x_1,x_2,x_3)$ is the spatial position, ψ is an eigenfunction, and λ is the corresponding eigenvalue. (Although the preceding formulation is for a single field, it can be generalized to multiple, correlated fields using tensors.[62])

The following is a step-by-step description of the application of GRPM:

(1) Apply a Fourier transform to the eigenvalue problem with respect to the homogeneous coordinates only. Because the integral in Eq. (A1) is a convolution with respect to the homogeneous coordinates, it reduces to a simple multiplication after transformation, and we are left with an integration over the inhomogeneous coordinates.[62] The result is

$$\int \hat{R}(\mathbf{k};\xi,\xi')\,\hat{\psi}(\mathbf{k};\xi')\,d\xi' = \lambda(\mathbf{k})\,\hat{\psi}(\mathbf{k};\xi). \tag{A2}$$

In the above, functions that have been transformed with respect to the homogeneous coordinates are indicated by a hat, e.g., $\hat{\psi}$. The inhomogeneous coordinates are indicated by ξ, and the wave numbers corresponding to the homogeneous coordinates are indicated by $\mathbf{k}$.

(2) Find the eigenfunctions and eigenvalues in the transformed domain. Let us designate the solutions as $\hat{\psi}_n(\mathbf{k};\xi)$, $\lambda_n(\mathbf{k})$, where $n=1,2,3,\ldots$, and the set is arranged in order of decreasing eigenvalue. The eigenfunctions are defined to be orthonormal such that

$$\int \hat{\psi}_m(\mathbf{k};\xi)\,\hat{\psi}_n^*(\mathbf{k};\xi)\,d\xi = \delta_{mn}, \tag{A3}$$

where $\delta_{mn}=1$ if $m=n$ and zero otherwise. (The asterisk indicates complex conjugation.)

(3) Generate the transformed synthetic field $\hat{s}(\mathbf{k};\xi)$ based on an eigenfunction series,

$$\hat{s}(\mathbf{k};\xi) = \sum_n a_n(\mathbf{k})\,\hat{\psi}_n(\mathbf{k};\xi), \tag{A4}$$

with the random expansion coefficients $a_n(\mathbf{k})$ set to

$$a_n(\mathbf{k}) = \sqrt{\lambda_n(\mathbf{k})}\exp[i\gamma_n(\mathbf{k})], \tag{A5}$$

in which $\gamma_n(\mathbf{k})$ is a random phase [a random number in the domain $[0,2\pi)$].

(4) Inverse Fourier transform $\hat{s}(\mathbf{k};\xi)$ to find $s(\mathbf{x})$.

The key to the procedure is Eq. (A5). By selecting the expansion coefficients in this manner, we ensure that the correlation function of the synthetic field matches the model correlation function; i.e.,

$$\hat{R}(\mathbf{k};\xi,\xi') = \langle \hat{s}(\mathbf{k};\xi)\hat{s}^*(\mathbf{k};\xi')\rangle. \tag{A6}$$

To demonstrate this, first note that by substituting Eq. (A4) into (A6), one has

$$\hat{R}(\mathbf{k};\xi,\xi') = \sum_m \sum_n \langle a_m(\mathbf{k}) a_n^*(\mathbf{k}) \rangle \hat{\psi}_m(\mathbf{k};\xi) \hat{\psi}_n^*(\mathbf{k};\xi').$$

Multiplying both sides by $\hat{\psi}_j(\mathbf{k};\xi')$, integrating with respect to ξ', and invoking orthonormality of the eigenfunctions, this becomes

$$\int \hat{R}(\mathbf{k};\xi,\xi') \hat{\psi}_j(\mathbf{k};\xi') d\xi' = \sum_m \langle a_m(\mathbf{k}) a_j^*(\mathbf{k}) \rangle \hat{\psi}_m(\mathbf{k};\xi).$$

Comparison with Eq. (A2) now yields

$$\sum_m \langle a_m(\mathbf{k}) a_j(\mathbf{k}) \rangle \hat{\psi}_m(\mathbf{k};\xi) = \lambda_j(\mathbf{k}) \hat{\psi}_j(\mathbf{k};\xi).$$

Finally, multiplying by $\hat{\psi}_n^*(\mathbf{k};\xi)$, integrating, and invoking orthonormality one more time produces

$$\langle a_m(\mathbf{k}) a_n^*(\mathbf{k}) \rangle = \delta_{mn} \lambda_n(\mathbf{k}). \tag{A7}$$

This equation shows that the EOF expansion coefficients are mutually uncorrelated and that the expected value of the magnitude squared of each coefficient equals the corresponding eigenvalue. The simplest way to choose the coefficients while satisfying this condition is Eq. (A5).

Application of GRPM to homogeneous turbulence is particularly simple and equivalent to RPM. In this case, the Fourier transform is applied to all coordinates, and the transformed eigenvalue problem corresponding to Eq. (A2) is simply

$$\Phi(\mathbf{k}) \hat{\psi}(\mathbf{k}) = \lambda(\mathbf{k}) \hat{\psi}(\mathbf{k}). \tag{A8}$$

The eigenfunctions are made orthonormal by setting $\hat{\psi}(\mathbf{k}) = 1$. Hence, $\Phi(\mathbf{k}) = \lambda(\mathbf{k})$ and (A5) becomes

$$a(\mathbf{k}) = \sqrt{\Phi(\mathbf{k})} \exp[i\gamma(\mathbf{k})]. \tag{A9}$$

The synthesized field is the inverse transform of $\hat{s}(\mathbf{k}) = a(\mathbf{k})$.

An inhomogeneous case of particular interest, since it applies to numerical calculations of scattering into an acoustic shadow zone, is the synthesis of turbulence in a vertical plane where the statistics depend on height. Designating the coordinates in the plane of interest as (x,z) (where x is horizontal), we need to determine the eigenfunctions of $R(x'-x,z,z') = \langle u(x,y,z) u(x',y,z') \rangle$. Since the turbulence is assumed homogeneous in x, we transform the correlation function with respect to the separation $x'-x$ prior to solving for the eigenfunctions, resulting in

$$\int_0^H F_\parallel(k_x,z,z') \hat{\psi}(k_x,z') dz' = \lambda(k_x) \hat{\psi}(k_x,z),$$

where

$$F_\parallel(k_x,z,z') = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(x,z,z') \exp(-ik_x x) dx \tag{A10}$$

is the 1-D cross spectrum (a spectrum with respect to x, and a correlation function with respect to z) and H is the domain height. For computational implementations the eigenvalue problem is solved by discretizing the domain in the z direction into N levels, and then numerically performing an eigenanalysis on the resulting matrix equation.[62,63] A set of N eigenfunctions $\hat{\psi}_n(k_x,z)$ and N eigenvalues $\lambda_n(k_x)$ results. The random expansion coefficients are

$$a_n(k_x) = \sqrt{\lambda_n(k_x)} \exp[i\gamma_n(k_x)]. \tag{A11}$$

The 1-D cross spectrum for the von Kármán model can be found from Eqs. (6) and (27), and transforming them as indicated by Eq. (A10), with result

$$F_\parallel(k_x,z,z') = \frac{2\sigma^2 l}{\sqrt{\pi}\Gamma(1/3)} \left(\frac{\zeta/2}{1+k_x^2 l^2} \right)^{5/6} \times \left[K_{5/6}(\zeta) - \frac{\zeta}{2} K_{1/6}(\zeta) \right], \tag{A12}$$

where $\zeta = (|z'-z|/l)\sqrt{1+k_x^2 l^2}$. [Note that $F_\parallel(k_x,z,z')$ reduces to Eq. (28) when $z=z'$.] For inhomogeneous turbulence, we take l to be a function of $\bar{z} = (z+z')/2$. Some ex-

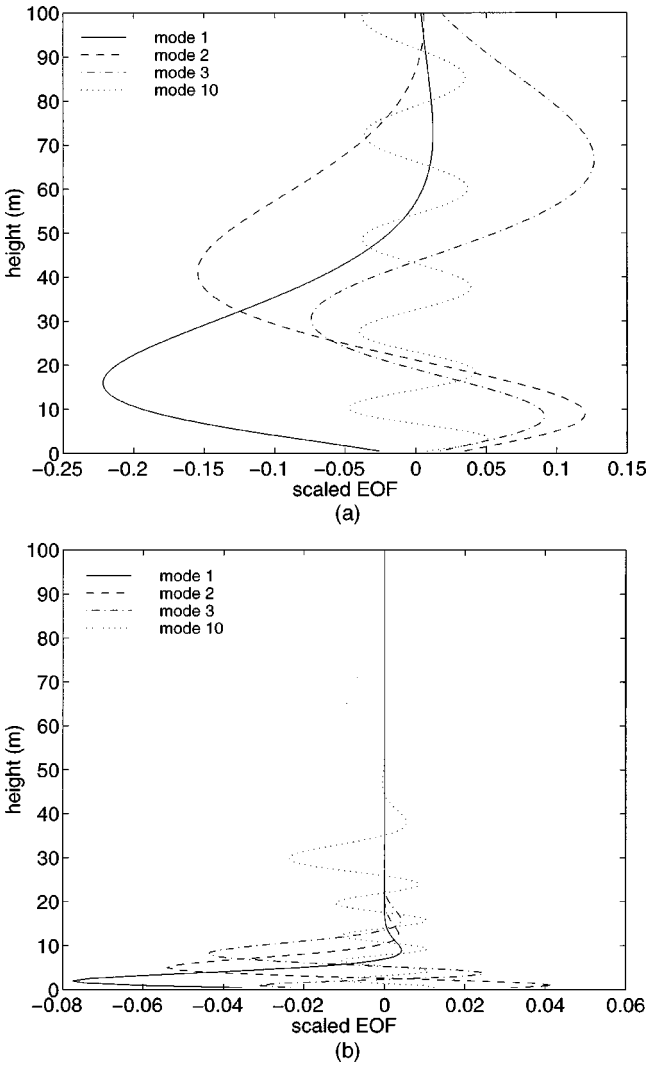


FIG. A1. Examples of empirical orthogonal functions (EOFs) for the von Kármán turbulence model as applied to shear turbulence. (''Mode'' is the EOF index n in the text.) The model length scale parameter is proportional to height, although the variance is independent of height. (a) Horizontal wavelength of 100 m. (b) Horizontal wavelength of 10 m. Abscissas are the amplitude of the EOF multiplied by the square root of the corresponding eigenvalue.

ample EOFs for shear turbulence, where $l = 1.8\overline{z}$ [Eq. (30)], are shown in Fig. A1. These EOFs, along with the others in the set, were used to synthesize the field shown in Fig. 10 as well as the shear-turbulence contribution in Fig. 11. The appearance of the EOFs is roughly similar to harmonic functions except that they are attenuated with increasing height.

[1] J. C. Wyngaard, ''Atmospheric Turbulence,'' Annu. Rev. Fluid Mech. **24**, 205–233 (1992).
[2] J. C. Kaimal and J. J. Finnegan, *Atmospheric Boundary Layer Flows: Their Structure and Measurement* (Oxford U.P., New York, 1994).
[3] D. Juvé, P. Blanc-Benon, and P. Chevret, ''Sound propagation through a turbulent atmosphere: influence of the turbulence model,'' in *Proceedings of the Sixth International Symposium on Long Range Sound Propagation*, edited by D. I. Havelock and M. R. Stinson (NRC Canada, Ottawa, Canada, 1994), pp. 270–282.
[4] V. E. Ostashev, ''Sound propagation and scattering in media with random inhomogeneities of sound speed, density and medium velocity,'' Waves Random Media **4**, 403–428 (1994).
[5] D. K. Wilson and D. W. Thomson, ''Acoustic propagation through anisotropic, surface-layer turbulence,'' J. Acoust. Soc. Am. **96**, 1080–1095 (1994).
[6] P. Boulanger, R. Raspet, and H. E. Bass, ''Sonic boom propagation through a realistic turbulent atmosphere,'' J. Acoust. Soc. Am. **98**, 3412–3417 (1995).
[7] M. Stinson and G. Daigle, ''Meteorological measurements for use in sound propagation calculations,'' in *Proceedings of the Seventh International Symposium on Long Range Sound Propagation* (Ecole Centrale de Lyon, Lyon, France, 1996), pp. 137–147.
[8] K. E. Gilbert, X. Di, and R. R. Korte, ''Distorted-wave Born approximation analysis of sound levels in a refractive shadow zone,'' in *Seventh International Symposium on Long Range Sound Propagation* (Ecole Centrale de Lyon, Lyon, France, 1996), pp. 373–389.
[9] V. E. Ostashev, B. Brähler, V. Mellert, and G. H. Goedecke, ''Coherence functions of plane and spherical waves in a turbulent medium with the von Karman spectrum of medium inhomogeneities,'' J. Acoust. Soc. Am. **104**, 727–737 (1998).
[10] D. K. Wilson, J. G. Brasseur, and K. E. Gilbert, ''Acoustic scattering and the spectrum of atmospheric turbulence,'' J. Acoust. Soc. Am. **105**, 30–34 (1999).
[11] H. A. Panofsky and J. A. Dutton, *Atmospheric Turbulence: Models and Methods for Engineering Applications* (Wiley, New York, 1984).
[12] L. Kristensen, D. H. Lenschow, P. Kirkegaard, and M. Courtney, ''The spectral velocity tensor for homogeneous boundary-layer turbulence,'' Boundary-Layer Meteorol. **47**, 149–193 (1989).
[13] J. Mann, ''The spatial structure of neutral atmospheric surface layer turbulence,'' J. Fluid Mech. **273**, 141–168 (1994).
[14] L. J. Peltier, J. C. Wyngaard, S. Khanna, and J. G. Brasseur, ''Spectra in the unstable surface layer,'' J. Atmos. Sci. **53**, 49–61 (1996).
[15] D. K. Wilson, ''Calculated coherence and extinction of sound waves propagating through anisotropic, shear-induced turbulent velocity fluctuations,'' J. Acoust. Soc. Am. **105**, 658–671 (1999).
[16] J. C. Wyngaard and L. J. Peltier, ''Experimental micrometeorology in an era of turbulence simulation,'' Boundary-Layer Meteorol. **78**, 71–86 (1996).
[17] V. N. Karavainikov, ''Fluctuations of amplitude and phase in a spherical wave,'' Akust. Zh. **3**, 175–186 (1957).
[18] L. A. Chernov, *Wave Propagation in Random Media* (McGraw-Hill, New York, 1960).
[19] A. N. Kolmogorov, ''The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers,'' C. R. Acad. Sci. URSS **30**, 301–305 (1941).
[20] V. I. Tatarskii, *The Effects of the Turbulent Atmosphere on Wave Propagation* (Keter, Jerusalem, 1971).
[21] A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, New York, 1978).
[22] V. E. Ostashev, *Acoustics in Moving Inhomogeneous Media* (E & FN Spon, London, 1997).
[23] J. C. Wyngaard, Y. Izumi, and S. A. Collins, ''Behavior of the refractive-index-structure parameter near the ground,'' J. Opt. Soc. Am. **61**, 1646–1650 (1971).
[24] G. A. Daigle, J. E. Piercy, and T. F. W. Embleton, ''Line-of-sight propagation through atmospheric turbulence near the ground,'' J. Acoust. Soc. Am. **74**, 1505–1513 (1983).
[25] M. A. Johnson, R. Raspet, and M. T. Bobak, ''A turbulence model for sound propagation from an elevated source above ground level,'' J. Acoust. Soc. Am. **81**, 638–646 (1987).
[26] B. Brähler, ''Zur Schallstreuung in turbulenten Fluiden (On Sound Scattering in Turbulent Fluids),'' Ph.D. thesis, Carl von Ossietzky Universität Oldenburg, 1996.
[27] T. von Kármán, ''Progress in the statistical theory of turbulence,'' Proc. Natl. Acad. Sci. U.S.A. **34**, 530–539 (1948).
[28] J. W. Strohbehn, ''Line-of-sight propagation through the turbulent atmosphere,'' Proc. IEEE **56**, 1301–1318 (1968).
[29] S. M. Flatté, R. Dahsen, W. H. Munk, K. M. Watson, and F. Zachariasen, *Sound Transmission Through a Fluctuating Ocean* (Cambridge U. P., Cambridge, England, 1979).
[30] V. E. Ostashev and D. K. Wilson, ''Relative contributions from temperature and wind velocity fluctuations to the statistical moments of a sound field in a turbulent atmosphere,'' Acustica **86**, 260–268 (2000).
[31] Some authors, including Flatté *et al.*, define the integral length scale by integrating the correlation function from $-\infty$ to ∞. The single-sided definition used in Eq. (2) is more customary in the turbulence literature.
[32] In Flatté *et al.*'s book (Ref. 29), a distinction is made between horizontal and vertical correlation lengths. In turbulence, however, the correlation length is independent of direction for small separations. This is the *local isotropy* property of turbulence discussed by Tatarskii (Ref. 20).
[33] For the purposes of this paper, it is assumed that the separation r in Eq. (3) lies within the inertial subrange, $\eta \ll r \ll L$, where η is the Kolmogorov microscale. If $r \ll \eta$, the power-law exponent would be $p = 3$ and L in this case is called the Taylor microscale. At the other extreme, the correlation function expansion is useful only when the Fresnel radius is much smaller than L, allowing derivation of Eq. (4).
[34] G. K. Batchelor, *The Theory of Homogeneous Turbulence* (Cambridge U.P., Cambridge, Great Britain, 1953).
[35] J. C. Kaimal, J. C. Wyngaard, Y. Izumi, and O. R. Coté, ''Spectral characteristics of surface layer turbulence,'' Q. J. R. Meteorol. Soc. **98**, 563–589 (1972).
[36] U. Högström, ''Review of some basic characteristics of the atmospheric surface layer,'' Boundary-Layer Meteorol. **78**, 215–246 (1996).
[37] H. Tennekes and J. L. Lumley, *A First Course in Turbulence* (MIT, Cambridge, MA, 1972).
[38] R. B. Stull, *An Introduction to Boundary Layer Meteorology* (Kluwer, Dordrecht, Germany, 1988).
[39] J. W. Deardorff, ''A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers,'' J. Fluid Mech. **41**, 453–480 (1970).
[40] S. J. Caughey and S. G. Palmer, ''Some aspects of turbulence structure through the depth of the convective boundary layer,'' Q. J. R. Meteorol. Soc. **105**, 811–827 (1979).
[41] D. K. Wilson, ''A three-dimensional correlation/spectral model for turbulent velocities in a convective boundary layer,'' Boundary-Layer Meteorol. **85**, 35–52 (1997).
[42] D. K. Wilson, ''On the application of turbulence spectral/correlation models to sound propagation in the atmosphere,'' in *Eighth International Symposium on Long Range Sound Propagation* (Pennsylvania State Univ., State College, PA, 1998), pp. 296–312.
[43] J. Højstrup, ''Velocity spectra in the unstable planetary boundary layer,'' J. Atmos. Sci. **39**, 2239–2248 (1982).
[44] A similar equation appears in Tatarskii's book (Ref. 20), except that a factor 2 replaces the $\frac{3}{2}$, and C_N^2 replaces C_V^2/c_0^2. The coefficient changes because the index-of-refraction field is scalar, whereas the velocity is vector. Ostashev (Ref. 4) has previously found similar modifications to scalar equations when they are adapted to vector fields.
[45] E. H. Brown and S. F. Clifford, ''On the attenuation of sound by turbulence,'' J. Acoust. Soc. Am. **60**, 788–794 (1976).
[46] J. O. Hinze, *Turbulence* (McGraw-Hill, New York, 1975).
[47] D. K. Wilson, ''Performance bounds for acoustic direction-of-arrival arrays operating in atmospheric turbulence,'' J. Acoust. Soc. Am. **103**, 1306–1319 (1998).
[48] A. L'Espérance, J. Nicholas, D. K. Wilson, D. W. Thomson, Y. Gabillet, and G. Daigle, ''Sound propagation in the atmospheric surface layer: Comparison of experiment with FFP predictions,'' Appl. Acoust. **40**, 325–346 (1993).
[49] R. Dashen, ''Path integrals for waves in random media,'' J. Math. Phys. **20**, 894–920 (1979).
[50] V. E. Ostashev, G. Goedecke, B. Brähler, V. Mellert, and H. Auvermann,

''Coherence functions of plane and spherical sound waves in the turbulent atmosphere with von Karman spectra of temperature and wind velocity fluctuations,'' in *Seventh International Symposium on Long Range Sound Propagation* (Ecole Centrale de Lyon, Lyon, France, 1996), pp. 349–357.
[51] F. M. Wiener and D. N. Keast, ''Experimental study of the propagation of sound over ground,'' J. Acoust. Soc. Am. **31**, 724–733 (1959).
[52] J. E. Piercy, T. F. W. Embleton, and L. C. Sutherland, ''Review of noise propagation in the atmosphere,'' J. Acoust. Soc. Am. **61**, 1403–1418 (1977).
[53] K. E. Gilbert, R. Raspet, and X. Di, ''Calculation of turbulence effects in an upward refracting atmosphere,'' J. Acoust. Soc. Am. **87**, 2428–2437 (1990).
[54] M. West, K. Gilbert, and R. A. Sack, ''A tutorial on the parabolic equation (PE) model used for long range sound propagation in the atmosphere,'' Appl. Acoust. **37**, 31–49 (1992).
[55] M. R. Stinson, D. I. Havelock, and G. A. Daigle, ''Simulation of scattering by turbulence into a shadow region using the GF-PE method,'' in *Proceedings of the Sixth International Symposium on Long Range Sound Propagation*, edited by D. I. Havelock and M. R. Stinson (NRC Canada, Ottawa, Canada, 1994), pp. 283–295.
[56] G. A. Daigle, ''Effects of atmospheric turbulence on the interference of sound waves above a finite impedance boundary,'' J. Acoust. Soc. Am. **65**, 45–49 (1979).
[57] D. Etling and R. A. Brown, ''Roll vortices in the planetary boundary layer: A review,'' Boundary-Layer Meteorol. **65**, 215–248 (1993).
[58] P. Chevret, P. Blanc-Benon, and D. Juvé, ''A numerical model for sound propagation through a turbulent atmosphere near the ground,'' J. Acoust. Soc. Am. **100**, 3587–3599 (1996).
[59] K. E. Gilbert, X. Di, S. Khanna, M. J. Otte, and J. C. Wyngaard, ''Electromagnetic wave propagation through simulated atmospheric refractivity fields,'' Radio Sci. **34**, 1413–1435 (1999).
[60] D. K. Wilson, J. C. Wyngaard, and D. I. Havelock, ''The effect of turbulent intermittency on scattering into an acoustic shadow zone,'' J. Acoust. Soc. Am. **99**, 3393–3340 (1996).
[61] J. L. Lumley, *Stochastic Tools in Turbulence* (Academic, New York, 1971).
[62] D. K. Wilson, ''Empirical orthogonal function analysis of the weakly convective atmospheric boundary layer. Part I: Eddy structures,'' J. Atmos. Sci. **53**, 801–823 (1996).
[63] P. Moin and R. D. Moser, ''Characteristic-eddy decomposition of turbulence in a channel,'' J. Fluid Mech. **200**, 471–509 (1989).

Evaluation of sound propagation models used in bottom volume scattering studies

Dan Li
Massachusetts Institute of Technology/Woods Hole Oceanographic Institution Joint Program in Oceanography/Oceanographic Engineering, Woods Hole, Massachusetts 02543

Dajun Tang
Applied Physics Laboratory, University of Washington, Seattle, Washington 98105

George V. Frisk
Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543

(Received 2 October 1998; revised 11 April 2000; accepted 14 June 2000)

The proper evaluation of sound propagation between sources/receivers and scatterers is important in characterizing bottom volume scattering. In this article, several sound propagation models used in bottom volume scattering studies are evaluated and their results compared to the exact solution obtained through a numerical wave number integration technique. It is found that Hines' approach [J. Acoust. Soc. Am. **88**, 324–334 (1990)] works well for the two isovelocity half-space case except when the grazing angle is close to the critical angle. The far-field approximation, given by Ivakin [Sov. Phys. Acoust. **32**(6), 492–496 (1986)] and Mourad and Jackson [J. Acoust. Soc. Am. **94**, 344–358 (1993)], has a performance depending upon the sound speed structure in the sediment. For an isovelocity slow bottom, it agrees well with the exact solution. However, discrepancies arise for an isovelocity fast bottom or a bottom with a complex sound speed structure. In addition, the appropriateness of using the equivalent surface scattering strength as a function of grazing angle in volume scattering characterizations is studied. In conclusion, precautions need to be taken in modeling both the propagation effects and the scattering mechanisms associated with the bottom volume scattering process. © *2000 Acoustical Society of America.* [S0001-4966(00)03409-3]

PACS numbers: 43.30.Gv, 43.30.Ft, 43.30.Hw [DLB]

I. INTRODUCTION

Acoustic wave scattering from the ocean bottom generally includes rough water/bottom interface scattering and subbottom volume scattering. While rough surface scattering has been the focus for many years, recent evidence shows that volume scattering, due to inhomogeneities and/or scattering layers within the sediment, could contribute significantly to bottom reverberation, especially when the seafloor is relatively flat and/or the the bottom attenuation is small. In his early work on the measurement of the bottom backscattering strength, Merklinger[1] speculated that subbottom inhomogeneities may be primary contributors to the backscattered field. Jackson *et al.*[2] and Lyons *et al.*[3] also recognized the importance of volumetric inhomogeneities in the bottom scattering. Tang *et al.*'s[4] analysis of high-frequency scattering data showed that the gas voids in the sediment can be significant scatterers. Tang *et al.*[5] also found random layers in a deep sea sediment to be the major cause of volume scattering at low frequencies.

The conventional quantities used to characterize the above two processes are the surface scattering coefficient and the volume scattering coefficient,[6] which historically are defined in the context of a plane wave being scattered by scatterers confined in an otherwise homogeneous medium. In ocean bottom scattering, however, the incident field is normally not a plane wave. The multipath and refractive effects resulting from the bottom sound speed structure complicate the scattering process and make it extremely difficult to pinpoint the scattering element in the classical sense. This problem will not diminish when the bottom volume scattering process is treated by an equivalent surface scattering process with the parameter of equivalent surface scattering coefficient being used. Therefore, the credibility of the scattering strength estimated from experimental data by this means needs to be examined.

Bottom volume scattering modeling usually includes two components: the scattering part, which encompasses the appropriate scattering mechanisms, and the propagation part, which takes into account propagation from the source to the scatterer and from the scatterer to the receiver. We will concentrate on propagation models in this study, which is equivalent to evaluating the Green's function between the source/receiver and the scatterer. The necessity to include poro-elastic effects in the scattering model remains an open question,[7] and we will regard the bottom as an acoustic fluid. The free-space Green's function, together with the transmission coefficient at the water/bottom interface, has been used in some work.[8,9] Despite its simplicity, it cannot incorporate the interface wave contribution at a subcritical grazing angle for a fast bottom. Since we are interested only in the methods that are good for the entire angular regime, this method will not be discussed in this article. Ivakin[10] and Mourad and Jackson[11] considered the incident wave on the interface to be a plane wave, with the incident direction varying over the insonified region to match the true incident direction. For convenience, we call it the far-field approximation in this

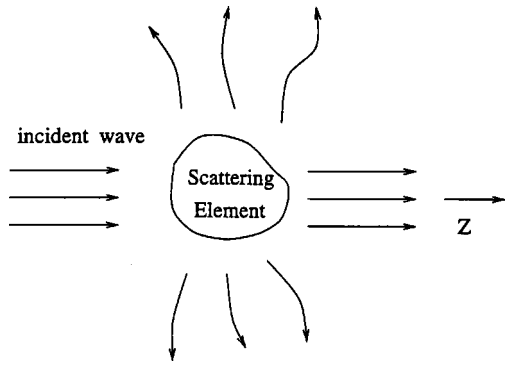


FIG. 1. Geometry of the classical scattering problem.

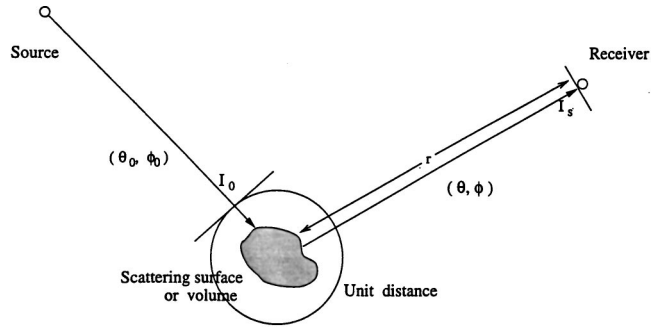


FIG. 2. Schematic illustration of the surface and volume scattering coefficient definitions.

article. This approach is actually a simplified version of the stationary phase method with only the most dominant stationary point contribution being included. Meanwhile, Hines[12] followed the stationary phase approach of Westwood[13] to find the transmitted field due to a point source in the two-isovelocity half-space case.

The goal of this article is to demonstrate the validity and limitations of the above-mentioned propagation models based on realistic geoacoustic parameters. Results from these models are compared with the exact solution obtained through a numerical wave number integration technique. Their influence on the estimation of the equivalent surface scattering strength is also studied in the article.

The paper is structured as follows: In Sec. II, we discuss the classical scattering problem and the definition of the conventional scattering cross section. The concept of an equivalent surface scattering strength which incorporates bottom volume scattering is also reviewed; in Sec. III, we describe the propagation models, and in Sec. IV, we compare the results of these models for some typical cases of interest.

II. THE SCATTERING CROSS SECTION

The conventional scattering cross section is defined when describing the scattering phenomenon associated with a single scatterer (e.g., a particle), as shown in Fig. 1. For simplicity, the incident plane wave is assumed to be propagating in the positive z direction with unit amplitude and suppressed time dependence $e^{-i\omega t}$:

$$\psi_i = e^{ikz}. \tag{1}$$

The scattered wave behaves as a spherical wave in the far field and is measured at infinity:

$$\psi_s = f(\theta,\phi)\frac{e^{ikr}}{r}, \quad \text{for } r\to\infty, \tag{2}$$

where r, θ, and ϕ are spherical coordinates with the origin at the scatterer. The scattering cross section is then defined as[14]

$$\sigma(\theta,\phi) = |f(\theta,\phi)|^2. \tag{3}$$

It is a measure of the scattered power in the (θ, ϕ) direction per unit solid angle, per unit incident intensity. Because the incident field is a plane wave and the observation point is at infinity, σ is a characterization of the scatterer independent of the source/receiver geometry.[15]

As stated before, in bottom scattering studies, scattering by the rough water/bottom interface and volume inhomogeneities within the bottom are two key processes. Any type of scattering modeling will inevitably encounter the problem of describing the scattering ability of a surface area and/or a volume. Figure 2 shows the concept underlying the definitions of a surface scattering coefficient σ_s and a volume scattering coefficient σ_v, which is similar to σ. Plane wave incidence is again assumed here, with the source and receiver in the far field. The incident intensity I_0 is measured at unit distance away from the scatterer. The scattered wave intensity I_s is obtained at the receiver and then multiplied by r^2 in order to compensate for spherical spreading of the scattered wave. Similar to the definition of σ, the scattering ability is characterized by the ratio of the scattered wave intensity to the incident intensity at unit distance from the scatterer per unit area or volume. It can be expressed in the following formulas:[6]

$$\sigma_s = I_s(\theta,\phi)r^2/[I_0(\theta_0,\phi_0)dS] \tag{4}$$

and

$$\sigma_v = I_s(\theta,\phi)r^2/[I_0(\theta_0,\phi_0)dV], \tag{5}$$

where (θ_0, ϕ_0) is the incident direction, (θ, ϕ) is the scattering direction, and dS and dV are the insonified area and volume, respectively.

Several points need to be noticed about the introduction of σ_s and σ_v. First, instead of a single scatterer, there exists a distribution of scatterers in the area S or volume V. Therefore, the total scattering area or volume is assumed to be divided into many differential elements, with each element scattered independently.[16] When the density of scatterers is sufficiently small or the correlation length of the surface/volume randomness is much smaller than the linear dimension of each scattering element, this assumption is reasonable. Nevertheless, the proper choice of the scattering element size needs to be stressed.

Second, in the definition of σ, a plane wave incident on the scatterer is assumed in order to exclude the effect of source position. Similarly for σ_s and σ_v, the incident wavefront is considered planar or locally planar, with the source in the far field. However, due to the possible multipath and refractive effects, this is generally not true for the bottom scattering problem. Third, for the same reasons, the scattered wave is not simply a spherical wave. As a result, the spread-

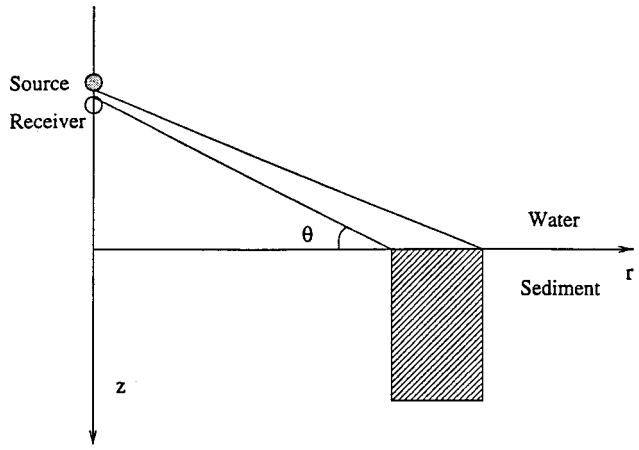


FIG. 3. Schematic illustration of the equivalent surface scattering coefficient definition.

ing loss at a distance r cannot be compensated for by the factor r^2, and in general, the phase is incorrectly calculated as well.

Despite the increasing interest in bottom volume scattering, it remains a difficult issue to characterize the scattering process. Unlike the rough surface scatterers which are concentrated at the water/bottom interface, the scatterers associated with volumetric scattering may be distributed over a large region. Scatterers may be located immediately beneath the water/bottom interface or a certain distance below the seafloor. As illustrated in Fig. 5, if the bottom sound speed is greater than that in the water and the scatterer is not at the interface but at depth Z, with β smaller than critical grazing angle, the scattered wave can reach the receiver through two paths *OAS* and *OBS*, therefore at two different angles. The path lengths are different so that the arrival times are not the same. As will be discussed later, the above multi-path effect will pose nontrivial problems for scattering strength estimations.

When the bottom attenuation is sufficiently high and the depth of acoustic penetration is much smaller than the distance between the source and the sediment, scattering from volume inhomogeneities within the bottom can be described by an equivalent surface scattering process.[11] As shown in Fig. 3, scattering from the shaded region, which is a slice of the scattering volume in an axisymmetric coordinate system, is attributed to scattering from the interface surface area of the volume. All of the scattering is assumed to be associated with the grazing angle θ because the difference would be minimal since the observation point (i.e., receiver) is far away. The equivalent surface scattering coefficient $\sigma_{\rm es}$, instead of the volume scattering coefficient σ_v, is then defined to quantify the scattering process. Mourad and Jackson[11] present a thorough description of this approach, with the result

$$\sigma_{\rm es} = \int \sigma_v |\Psi_1|^4 \, dz, \tag{6}$$

where

$$\Psi_1 = \Psi_v / \Psi_s . \tag{7}$$

Here z is the depth coordinate, Ψ_v and Ψ_s are the incident field as it would be measured at the volume element dV and the area element dA, respectively. The quantity $\sigma_{\rm es}$ can be viewed as a measure of the scattering contribution from the whole shaded region while Ψ_1 is the normalized field as if the scatterers were at the surface.

The concept of an equivalent surface scattering coefficient has been used extensively in scattering data analysis. In the high-frequency, high-attenuation case, the incident wave may penetrate to relatively shallow depths. However, at low frequencies and attenuations, the ensonification can encompass a substantial portion of the sediment column. The sediment layering and the sound speed gradient cause the scatterers to be ensonified by an incident field consisting of refracted, multiple reflected paths. Even if we neglect the multiple scattering effect, which is to ignore the rescattering of the scattered wave, it is still very difficult to relate the scattered returns recorded in an experimental time series to the corresponding scatterers. As can be seen in Fig. 3, the thickness of the shaded region would be very large for deep penetration. In the above calculation of the equivalent surface scattering strength, all the contributions from each shaded ''vertical bar'' would be ascribed to only one scattering angle, which can be justified for shallow penetration in a high-frequency situation but is invalid in the low-frequency case. In fact, the volume element could affect the scattering at different angles due to multipath effects. As a result, the angular dependence of the predicted scattering strength might be error-prone. In a shallow water environment, the limited water depth would further deteriorate the problem, since the distance between the source/receiver and the water/bottom interface is confined. All of the above would directly affect the effectiveness of using the scattering strength as a criterion in a model/data comparison.

III. BOTTOM PROPAGATION MODELS

Propagation, i.e., the Green's function, plays an important role in the modeling of the scattering process. In this section, three methods will be described for calculating the propagation from a point source in the water column to a scatterer in the sediment. The propagation from the scatterer to a receiver in the water can be obtained accordingly by the principle of reciprocity.

The wave equation in an inhomogeneous medium[17] can be written as

$$\rho(\mathbf{r}) \nabla \cdot \left(\frac{1}{\rho(\mathbf{r})} \nabla p(\mathbf{r},t) \right) - \frac{1}{c^2(\mathbf{r})} \frac{\partial^2 p(\mathbf{r},t)}{\partial t^2} = f(\mathbf{r},t), \tag{8}$$

where $p(\mathbf{r},t)$ represents the pressure, $\rho(\mathbf{r})$ represents the density, $c(\mathbf{r})$ represents the sound speed, and $f(\mathbf{r},t)$ represents the source term as a function of space $\mathbf{r}$ and time t. If constant density and harmonic time dependence $[\exp(-i\omega t)]$ are assumed, Eq. (8) becomes the Helmholtz equation

$$[\nabla^2 + k_0^2(\mathbf{r})(1+\epsilon(\mathbf{r}))] P(\mathbf{r},\omega) = F(\mathbf{r},\omega), \tag{9}$$

where $k_0(\mathbf{r})$ is the medium wave number for the background sound speed $c_0(\mathbf{r})$ and frequency ω, with

$$k_0(\mathbf{r}) = \frac{\omega}{c_0(\mathbf{r})}, \tag{10}$$

and

$$f(\mathbf{r},t) = F(\mathbf{r},\omega)e^{-i\omega t}, \quad p(\mathbf{r},t) = P(\mathbf{r},\omega)e^{-i\omega t}.$$

The parameter $\epsilon(\mathbf{r})$ is the perturbation to the wave number due to the inhomogeneous sound speed. Using a perturbation method, the total field $P(\mathbf{r})$ (for simplicity, ω will be implicit from now on) can be expressed as the summation of the incident field $P_0(\mathbf{r})$ and the scattered field $P_s(\mathbf{r})$:

$$P(\mathbf{r}) = P_0(\mathbf{r}) + P_s(\mathbf{r}), \tag{11}$$

where the incident field satisfies

$$[\nabla^2 + k_0^2(\mathbf{r})]P_0(\mathbf{r}) = F(\mathbf{r}). \tag{12}$$

We assume a point source at $\mathbf{r} = \mathbf{r_s}$ and the Green's function can be obtained from

$$[\nabla^2 + k_0^2(\mathbf{r})]G(\mathbf{r},\mathbf{r_s}) = \delta(\mathbf{r} - \mathbf{r_s}) \tag{13}$$

and the boundary conditions. The scattered field P_s satisfies

$$[\nabla^2 + k_0^2(\mathbf{r})]P_s(\mathbf{r}) = -k_0^2(\mathbf{r})\epsilon(\mathbf{r})P(\mathbf{r}). \tag{14}$$

The integral equation for the solution of the scattered field is

$$P_s(\mathbf{r}) = -k_0^2(\mathbf{r}) \int G(\mathbf{r},\mathbf{r}')\epsilon(\mathbf{r}')P(\mathbf{r}')\, d\mathbf{r}'. \tag{15}$$

From Eq. (15), it can be seen that the scattered field and the Green's function are closely related. The Green's function, which characterizes the propagation from the source to the scatterer and from the scatterer to the receiver, plays an important role in modeling of the bottom scattering process. In this section, three methods will be described for calculating the propagation from a point source in the water column to a scatterer in the sediment. The propagation from the scatterer to a receiver in the water can be obtained accordingly by the principle of reciprocity.

There are several ways to solve Eq. (13). For simplicity, we will consider a range-independent, horizontally stratified environment, where if we choose a cylindrical coordinate system, with the vertical z axis passing through the source and the r axis being parallel to the interfaces, it becomes an axisymmetric propagation problem. At this point, we will assume unit source strength. The Green's function can then be calculated by means of wave number integration. For the scenario shown in Fig. 4, we have the Green's function in the sediment $G(r,z)$ as follows,

$$G(r,z) = \int_0^\infty g(k_r,z)k_r J_0(k_r r)\, dk_r$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} g(k_r,z)k_r H_0^{(1)}(k_r r)\, dk_r, \tag{16}$$

where k_r is the horizontal wave number, J_0 is the zeroth-order Bessel function, $H_0^{(1)}$ is the zeroth-order Hankel function of the first kind, and $g(k_r,z)$ satisfies the equation

$$\left[\frac{d^2}{dz^2} + (k^2(z) - k_r^2)\right]g(k_r,z) = 0 \tag{17}$$

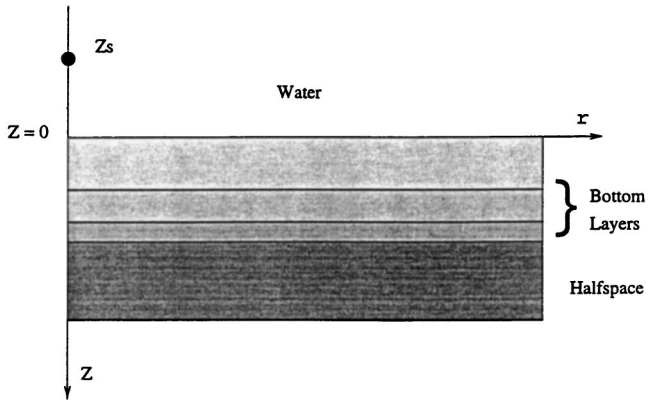


FIG. 4. The experimental scenario.

and the boundary conditions of continuity of normal particle velocity and pressure across each interface. When $k_r r \gg 1$, we have

$$G(r,z) = \frac{1}{\sqrt{2\pi r}} \int_{-\infty}^{\infty} g(k_r,z)\sqrt{k_r}e^{i(k_r r - \pi/4)}\, dk_r. \tag{18}$$

Equation (18) can be evaluated numerically. The procedure of implementation can be found in Schmidt's SAFARI manual[18] and Ref. 19. In fact, the numerical solution of the wave equation in Sec. IV is obtained primarily by using SAFARI. One of the advantages of SAFARI is that it generally provides stable results for complicated sediment structures.

A point source is located at $(0,z_s)$. If an isovelocity water column is assumed so that the water wave number $k_0 = \omega/c_0$, the depth-dependent Green's function in the water column will be

$$g_w(k_r,z) = -\frac{e^{ik_{z_0}|z-z_s|}}{4\pi i k_{z_0}}, \tag{19}$$

where $k_{z_0} = \sqrt{k_0^2 - k_r^2}$ is the vertical wave number in the water. By matching the boundary conditions at the water/sediment interface $z = 0$, it is clear that the depth-dependent Green's function in the sediment can be expressed in terms of $-e^{ik_{z_0}|z_s|}/4\pi i k_{z_0}$. In other words, we have

$$g(k_r,z) = -g_1(k_r,z)\frac{e^{ik_{z_0}|z_s|}}{4\pi i k_{z_0}}, \tag{20}$$

where $g_1(k_r,z)$ can be viewed as the solution to Eq. (17) with a plane wave incident on the water/sediment interface. Using the coordinates shown in Fig. 4, Eq. (18) now becomes

$$G(r,z) = \frac{e^{i\pi/4}}{\sqrt{2\pi r}} \int_{-\infty}^{\infty} \frac{g_1(k_r,z)\sqrt{k_r}e^{i(k_r r - k_{z_0} z_s)}}{4\pi k_{z_0}} dk_r. \tag{21}$$

The stationary phase method is often used to evaluate an integral of this type. Following Ivakin's approach,[10] one obtains

$$G(r,z) = \frac{e^{ik_0 R}}{4\pi R} g_1(k_0 \cos\beta, z), \tag{22}$$

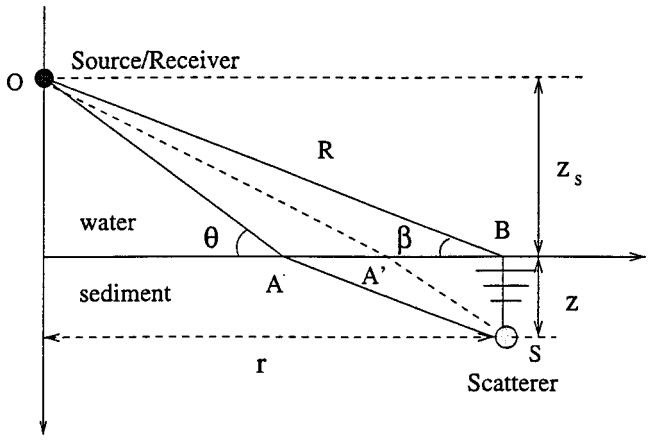


FIG. 5. Schematic illustration of scattering geometry for a two-isovelocity half-space case.

where $R=\sqrt{r^2+z_s^2}$ is the distance between the source and the point at the interface right above the scatterer in the sediment, and $\beta=\tan^{-1}|z_s|/r$ as shown in Fig. 5. For simplicity, a two half-space scenario is being considered. In the above approach, the propagation can be interpreted as a spherical wave traveling from the source to the interface point *B* above the scatterer and then a plane wave traveling from point *B* to the scatterer. Conventional asymptotic analysis for a slow bottom shows that there is only one eigenray path. Thus, in Ivakin's approach, path *OBS* is used to approximate path $OA'S$. Point *B* is the approximate stationary point and A' is the exact one. For a fast bottom, there are two cases. When the incident grazing angle β is greater than the critical angle, the situation is similar to that for a slow bottom. Only one eigenray path exists and *A* is the exact stationary point. However, when β is smaller than the critical angle, both Path *OAS* and *OBS* will exist. Paths *OAS* and *OBS* correspond to the refracted wave and the evanescent wave, respectively. If β is close to the critical angle, the ray interpretation will not be appropriate.[13] There are two potential errors in this approach. First, point *B* is not the same as point *A* or A', which will bring in a certain degree of error for β greater than the critical angle. Second, for β smaller than the critical grazing angle, the stationary point *A* contribution, i.e., the refracted wave, has been neglected. When the bottom attenuation is high and the scatterer is close to the water/sediment interface, this approach is justified, since the longer path of the refracted wave will cause it to be greatly attenuated. However, we will be able to see some differences if the above conditions cannot be satisfied. Mourad and Jackson's approach[11] is somewhat similar to Ivakin's. Instead of regarding the propagation as consisting of a spherical wave traveling from the source to the interface point, the incident wave is considered to be a plane or locally plane wave with incident direction allowed to vary over the insonified region in order to match the true incident direction. For convenience, we will call the above approach the far-field approximation in this article.

The formula that Hines derived in his paper,[12] also a far-field approach, works for a two-isovelocity half-space environment only. He applied the stationary phase method as well. In order to distinguish it from the above-mentioned far-field approximation, we will name it Hines' method in the later comparison. In a two-isovelocity half-space situation, $g_1(k_r,z)$ in Eq. (21) can be further expressed as

$$g_1(k_r,z)=T_{\rm ws}e^{ik_{z_1}z}, \tag{23}$$

where $T_{\rm ws}$ is the transmission coefficient from the water to the sediment, k_{z_1} is the vertical wave number in the sediment, and z is the distance from the interface to the scatterer. If Eq. (21) is rewritten in terms of the grazing angle $\theta_i=\cos^{-1}k_r/k_0$, we have

$$G(r,z)=e^{i\pi/4}\sqrt{\frac{k_0}{2\pi r}}\int_{i\infty-\pi/2}^{-i\infty+\pi/2}T_{\rm ws}(\theta_i)\sqrt{\cos\theta_i}\times e^{ik_0(r\cos\theta_i-z_s\sin\theta_i+z\sqrt{n_{\rm ws}^2-\cos^2\theta_i})}\,d\theta_i, \tag{24}$$

where $n_{\rm ws}=k_1/k_0$, and k_1 represents the medium wave number in the sediment. Equation (24) is a little different from that in Hines' paper since we choose θ_i to be the incident angle from the water to the sediment instead of the opposite. The phase of the integrand is now

$$f(\theta_i)=ik_0(r\cos\theta_i-z_s\sin\theta_i+z\sqrt{n_{\rm ws}^2-\cos^2\theta_i}), \tag{25}$$

where $n_{\rm ws}$ can be complex to account for the attenuation in the sediment. According to the stationary phase method, the stationary point can be found from the following:

$$\frac{df(\theta_i)}{d\theta_i}=-ik_0\left(r\cdot\sin\theta_i+z_s\cdot\cos\theta_i-\frac{z\cdot\sin(2\theta_i)}{2\sqrt{n_{\rm ws}^2-\cos^2\theta_i}}\right)=0. \tag{26}$$

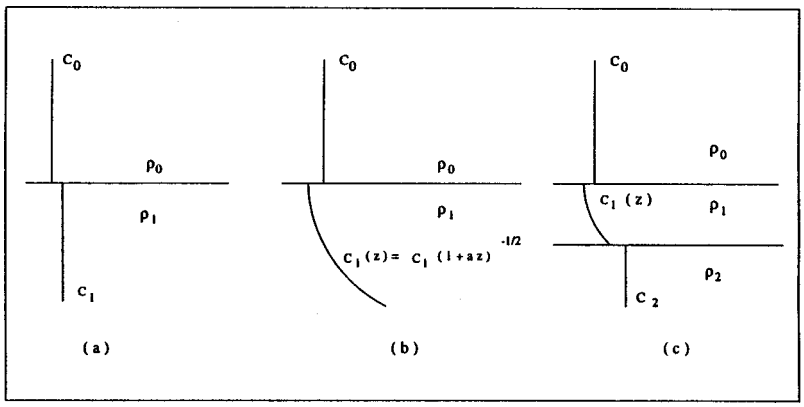


FIG. 6. The sound speed profile for three types of bottom: (a) two isovelocity half spaces, (b) two half spaces with a sound speed gradient in the bottom, and (c) three layers with a sound speed gradient in the middle layer.

There will be two solutions to Eq. (26) for a fast bottom. The solution corresponding to the evanescent wave can be obtained by perturbing the incident angle β, where

$$\beta = \tan^{-1}(-z_s/r). \tag{27}$$

The real solution can be found by performing a first-order Taylor expansion,

$$\theta_e = \beta - \frac{f'(\beta)}{f''(\beta)}, \tag{28}$$

where the prime ($'$) represents the derivative with respect to the incident angle. We have to use a complex root finder to obtain the second solution θ_r, which corresponds to the refracted wave. The perturbation method suggested by Hines has some difficulties when the actual stationary point is close to a singular point, the critical angle in this case. It might also be a problem for θ_e and is possibly a cause of the overestimate for the field in the region close to the critical angle as can be seen later.

For the two half-space, fast-bottom situation, the incident spherical wave is therefore decomposed into two plane waves: the refracted wave and the evanescent wave. The evanescent wave exists only in the subcritical grazing angle region. The stationary phase method gives the results

$$Q(\theta) = \frac{A(\theta)}{B(\theta)}, \tag{29}$$

where

$$A(\theta) = T_{\mathrm{ws}}(\theta)\sqrt{\cos\theta}\, e^{ik_0(r\cos\theta - z_s\sin\theta + z\sqrt{n_{\mathrm{ws}}^2 - \cos^2\theta})}, \tag{30}$$

and

$$B(\theta) = \sqrt{r\left(r\cos\theta - z_s\sin\theta - z\left[\frac{\cos(2\theta)}{\sqrt{n_{\mathrm{ws}}^2 - \cos^2\theta}} - \frac{\sin^2(2\theta)}{4\sqrt{(n_{\mathrm{ws}}^2 - \cos^2\theta)^3}}\right]\right)}. \tag{31}$$

For the refracted wave

$$G_r = Q(\theta_r), \tag{32}$$

and for the evanescent wave

$$G_e = Q(\theta_e). \tag{33}$$

From the results in the next section, this method can be seen to be reasonably good for the two isovelocity half-space case, except in the region close to the critical angle, where there are more than two stationary points, as pointed out by Westwood.[13] However, it is not trivial to find all the stationary points for a bottom with multiple layers or with a sound speed gradient because the phase term of the integrand in Eq. (19) becomes much more complicated.

IV. COMPARISON OF PROPAGATION MODELS

The far-field approximation and Hines' method have been implemented in bottom scattering modeling, especially in high-frequency problems. Yet their validity and limitations need to be recognized when dealing with general experimental scenarios. In this section, based on certain sets of realistic geoacoustic parameters, the above two propagation models will be evaluated through comparisons with exact solutions obtained by numerical wave number integrations. An effort will also be made to identify their possible influence on the quantification of bottom scattering.

We will concentrate on three types of fluid models (cf. Fig. 6): two isovelocity half-spaces, two half-spaces with an upward-refracting sound speed profile in the sediment, and a layer with an upward-refracting profile between two isovelocity half-spaces. Most of the geoacoustic parameters used in the comparisons are measurement results or extrapolations of measured data which have appeared in the literature. In this section, f will stand for frequency, and c_i and ρ_i will stand for the sound speed and density in the overlying water ($i=0$), just beneath the water/sediment interface ($i=1$), and in the lower half space ($i=2$) in Fig. 6(c). The density is chosen to be constant in each layer. The loss factor δ_i ($i=0,1,2$) is the ratio of the imaginary and real parts of the complex wave number at the top of each sediment layer. The quantity g represents the sound speed gradient, and H stands for the middle layer thickness in the three-layer case. The wavelength in the water is given by λ, z_s is the source height above the water/bottom interface, and z is the distance between the scatterer in the bottom and the water/sediment interface. The quantity z_s and z are scaled relative to wavelength λ. The source height z_s is chosen to be 100λ (theoretically it should be at infinity) in order to avoid the near-field effect. In each case, the transmission loss is plotted versus grazing angle β and is defined as the ratio in decibels between the acoustic pressure at the scatterer for a simple point source and the pressure produced at a distance of 1 m from the same source.

A. Two isovelocity half-spaces

The sound speed in the sediment depends on the sediment type. The sediment/water sound speed ratio could be larger or smaller than 1, which determines the existence of a critical angle. According to Hamilton,[20] for soft sediments such as silty clay, the sound speed just beneath the water/sediment interface is usually less than that of the overlying water. The geoacoustic parameters used in Fig. 7, which are consistent with Continental Rise data and taken from Fig. 3 in Mourad and Jackson's paper,[11] are used as an example. From the field at scatterers situated at $z=0$ and $z=-2\lambda$, it can be seen that results of the far-field approximation and Hines' method are almost identical to that of the numerical wave number integration. There is no critical angle and only

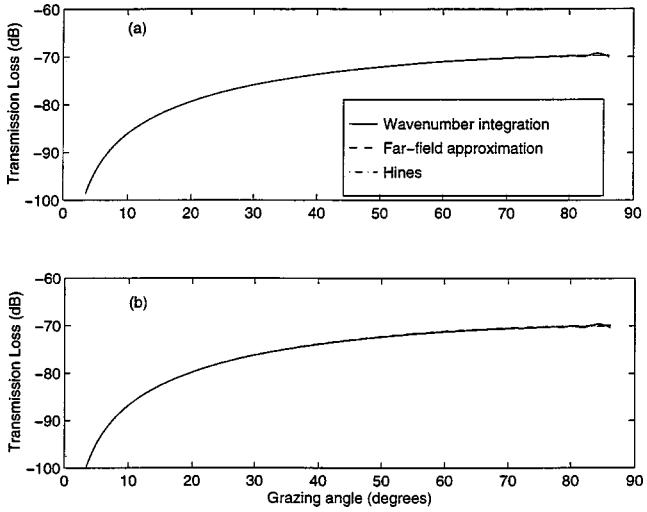


FIG. 7. Comparison of transmission loss in the environment with no critical angle. Parameters are $f=500$ Hz, $c_0=1530$ m/s, $c_1=1510$ m/s, $\rho_1=1.72$ g/cm^3, $\delta=0.001\,64$, and $z_s=100\lambda$. The transmitted fields at (a) $z=0$ and (b) $z=-2\lambda$.

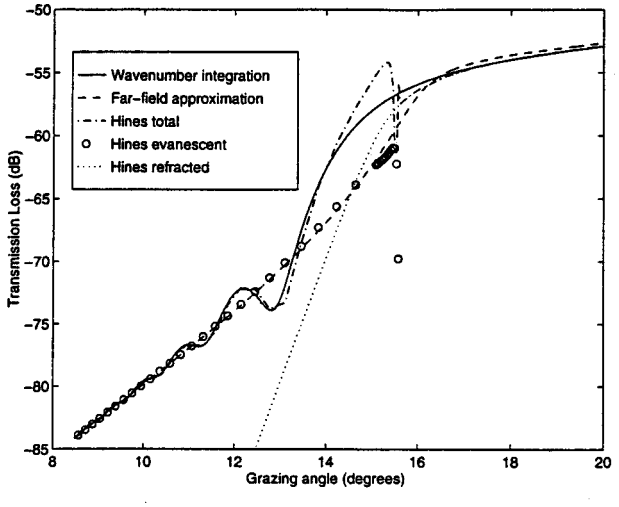


FIG. 9. The detailed comparison. Parameters are the same as those in Fig. 8(c).

one stationary point, i.e., scatterers are insonified by the refracted wave only. Both approaches are good approximations in this situation. One would expect the calculation of the equivalent surface scattering strength to be appropriate and accurate without including the multipath effect. Since z_s and z are both scaled relative to λ, the results obtained here and for the two isovelocity half-space case should be frequency independent if the bottom attenuation is linear with frequency, which is an open question.[21–23]

For sandy bottoms and turbidite bottoms in certain areas, the sediment/water sound speed ratio is larger than 1. Input parameters in Fig. 8 correspond to the Hudson Canyon experiment reported by Rogers *et al.*,[24] where the bottom is described as silty sand and silty clay. We choose the fre-

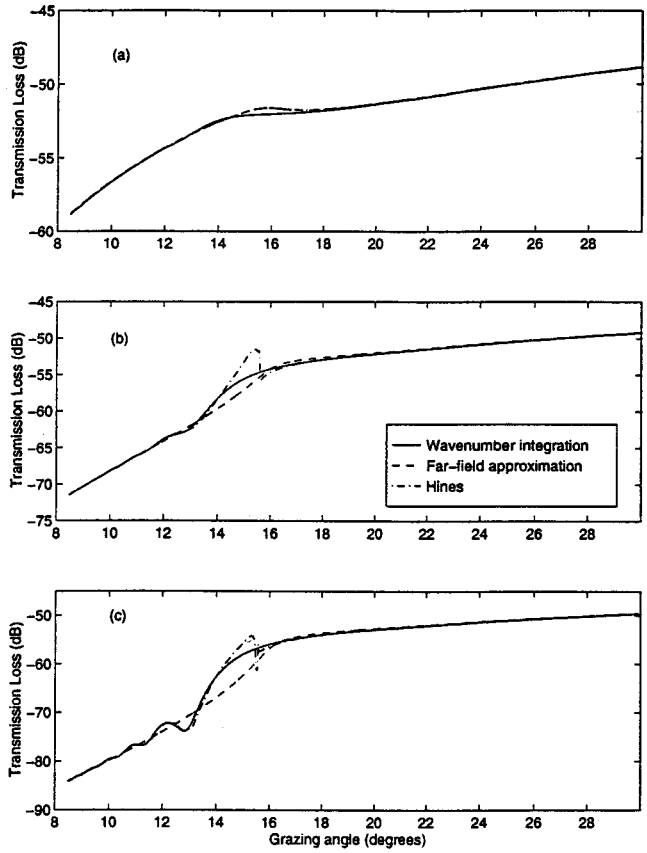


FIG. 8. Comparison of transmission loss with the critical angle effect for a low attenuation bottom. Parameters are $f=500$ Hz, $c_0=1500$ m/s, $c_1=1560$ m/s, $\rho_1=1.8$ g/cm^3, $\delta=0.0027$, and $z_s=100\lambda$. The transmitted fields at (a) $z=0$, (b) $z=-\lambda$, and (c) $z=-2\lambda$.

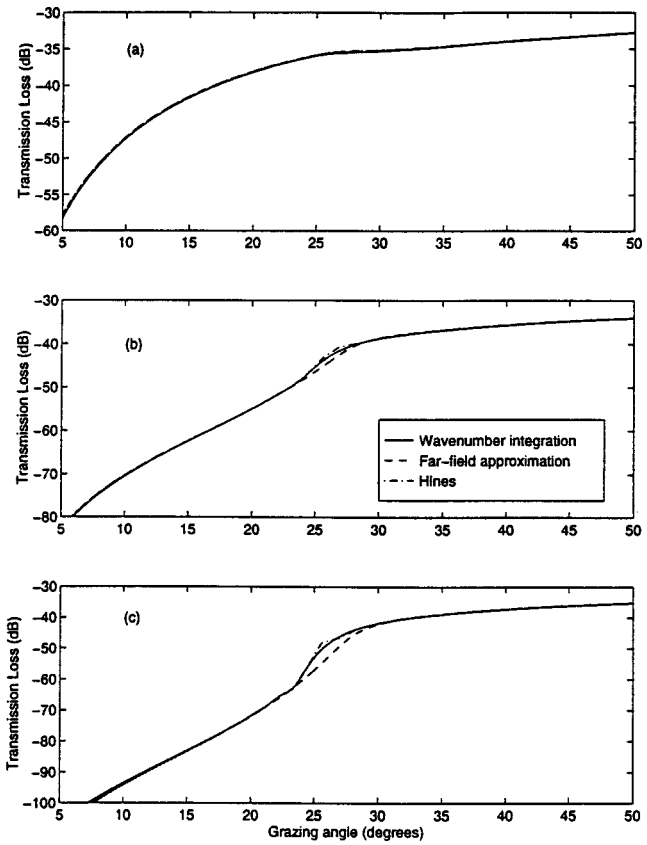


FIG. 10. The critical angle effect for a high-attenuation bottom. Parameters are $f=40\,000$ Hz, $c_0=1500$ m/s, $c_1=1689$ m/s, $\rho_1=1.97$ g/cm^3, $\delta=0.0166$, and $z_s=100\lambda$. The transmitted fields at (a) $z=0$, (b) $z=-\lambda$, and (c) $z=-2\lambda$.

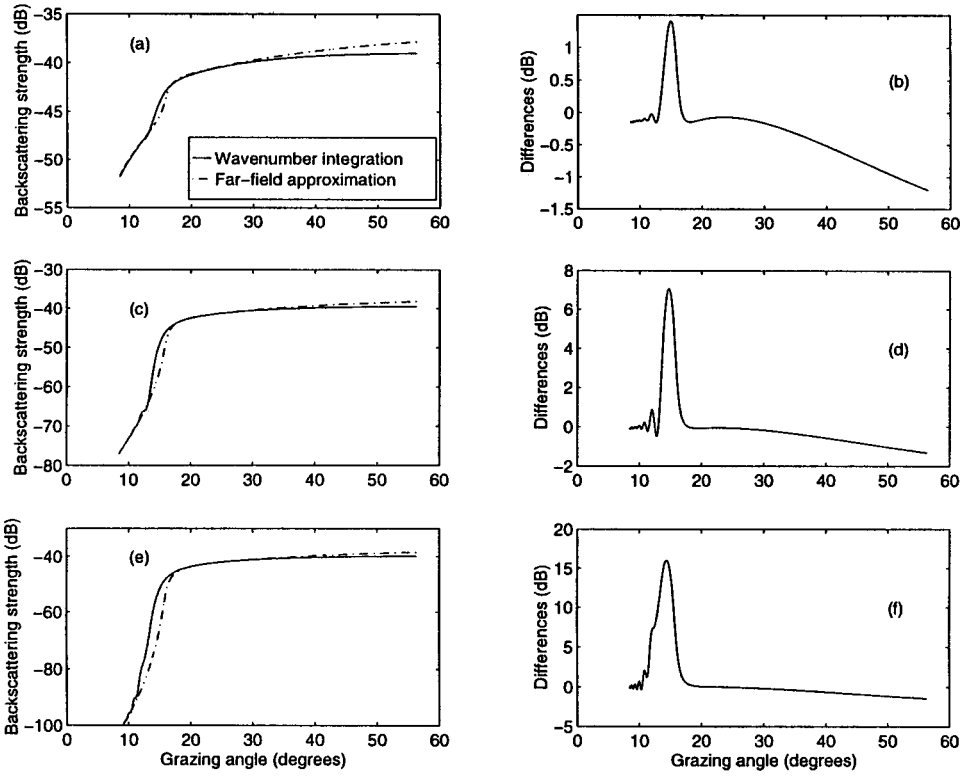


FIG. 11. The equivalent surface backscattering strength for a low-attenuation bottom assuming point scatterers. Parameters are the same as those in Fig. 8 and $\sigma_v = 0.000\,03$. The backscattering strength and the difference as the scattering layer begins at (a), (b) $z=0$, (c), (d) $z=-\lambda$, and (e), (f) $z=-2\lambda$.

quency to be 10 kHz and assign the bottom attenuation to be 0.1 dB/m/kHz ($\delta_i = 0.0027$), which is at the high end for this type of bottom. The transmitted field is again shown at different depths in Fig. 8. It is found that the agreement between Hines' method and the numerical wave number integration technique is fairly good except for some overestimate in the region just below the critical grazing angle [$\theta_c = \cos^{-1}(c_0/c_1)$] of 16 degrees and is indicated in the figure. The far-field approximation, on the other hand, tends to underestimate the field at subcritical grazing angles. Differences will increase with the depth of the scatterer (cf. Fig. 8), and with lower attenuation values. In Fig. 9, we add the evanescent wave and refracted wave contributions to Fig. 8(c). For the far-field approximation, it is clear that the neglect of the refracted wave in the subcritical angle region is the cause of the underestimate. The small error at greater than the critical angle can be attributed to the difference between the true stationary point and that used in the approximation. As for Hines' method, the match is excellent everywhere except in a small region at less than the critical angle. This is the transition area from a field composed of one stationary point to a field composed of two stationary points, where the approximation would normally be invalid, according to Westwood.[13] Another possible cause might be the failure of the first-order perturbation method to find the stationary point close to the singularity associated with the critical angle.

For a sandy bottom, the attenuation will be much higher than that for a soft bottom. The geoacoustic parameters in Fig. 10 are from the CBBL Panama City site.[7] Frequency of 40 kHz is chosen, the same as that in the experiment. It is a sandy bottom and the attenuation was found to be 0.524 dB/m/kHz, which is equivalent to $\delta = 0.0166$. The critical angle is now 27.4 degrees. The figure shows results similar to those in Fig. 8 for the field at three different depths. Nevertheless, the difference between the far-field approximation and the numerical wave number integration is smaller than that in the lower attenuation bottom. The interference between the refracted wave and the evanescent wave is hardly noticeable at subcritical angles. The explanation for this is that the refracted wave has a longer ray path than that of the evanescent wave and is therefore greatly attenuated at long ranges in the high-attenuation bottom. In other words, the evanescent wave contribution will prevail at subcritical angles. So we see from the above that the single stationary point treatment in the far-field approximation works better in a high-attenuation bottom than in a low-attenuation one.

The effect of different propagation models on the estimation of scattering strength is of considerable interest in bottom scattering modeling. In Figs. 11 and 12, we have calculated the equivalent surface scattering strength [cf. Eq. (6)] associated with volume backscattering using the same parameters as those in Figs. 8 and 10. Point scatterers are assumed here for simplicity, and the volume scattering cross section σ_v is chosen to be 0.000 03 in Fig. 11 and 0.000 906 in Fig. 12. Both in the ARSRP low-frequency sediment scattering experiment and the CBBL high-frequency scattering experiment at the Eckernfoerde Bay site, strong scattering layers are found to be a certain distance away from the water/sediment interface.[5,4] So, we will choose a scattering layer with a thickness of 50λ here starting from different depths below the interface. Figures 11 and 12(b), (d), and (f) show the difference in backscattering strength between the exact solution and the far-field approximation result. Since results from Hines' method fit exact solutions very well, but

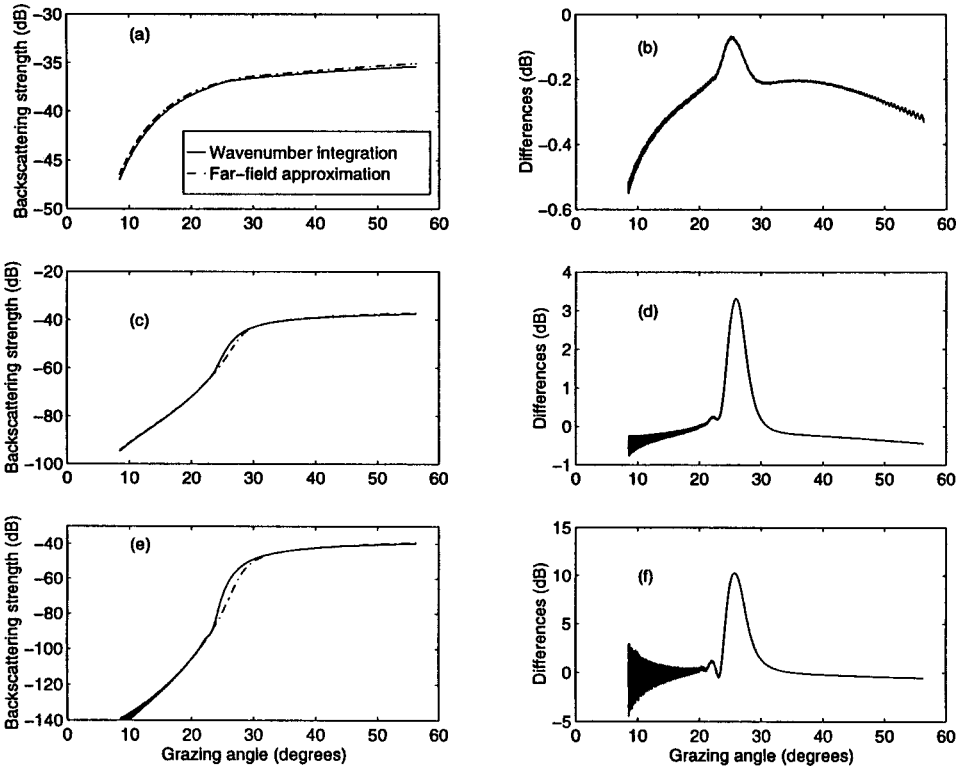


FIG. 12. The equivalent surface backscattering strength for a high-attenuation bottom assuming point scatterers. Parameters are the same as those in Fig. 10 and $\sigma_v = 0.000\,906$. The backscattering strength and the difference as the scattering layer begins at (a), (b) $z=0$, (c), (d) $z=-\lambda$, and (e), (f) $z=-2\lambda$.

in a very narrow region, we consider that it would yield similar outputs and will not discuss it here.

From Fig. 11 it is evident that the underestimate of the field at subcritical angles for the far-field approximation leads to differences in equivalent surface backscattering strength. However, the biggest difference is only about 1 dB when the scattering layer starts right at the interface. On the other hand, when the scattering layer begins at greater depths, the difference is much larger. For the scattering layer starting at two wavelengths beneath the interface, up to 15 dB difference appears. Also, we notice that the actual level of the backscattering strength in this case is only several dB lower than that for the scattering layer starting from the interface. Therefore, they can be important contributors to overall bottom scattering, which highlights the necessary precautions in scattering modeling using the far-field approximation in this environment. When the bottom attenuation is low, the penetration is deep. The multipath effect is inevitable. The scattered returns will reach the receiver via different paths, therefore different scattering angles, which makes the angular dependence of the scattering strength less meaningful. All of these results indicate that when the far-field approximation is applied to specific scenarios and the equivalent surface scattering strength is used to interpret data, one should examine the appropriateness of the modeling.

Figure 12 shows similar results for a sandy bottom. The difference in backscattering strength is smaller when compared with that in Fig. 11. Numerical errors can even be seen in Figs. 12(e) and (f). The scattering level for a scattering layer deep in the sediment is much lower than that at the interface. Since the sand is usually a strong scatterer and the bottom penetration is much shallower due to high attenuation, the problems that are addressed regarding Fig. 11 are much less serious. Both the far-field approximation and the equivalent surface scattering strength are appropriate.

B. Two half-spaces with sound speed gradient in the lower half

Sound speed gradients exist in both the water column and sediment in real ocean environments[25] and have significant effects on sound propagation. The gradient in the water column is generally small compared with that in the sediment and will be neglected. We will therefore focus on the effect of the sound speed gradient in the sediment. We will not include Hines' method in the comparisons from now on because it does not apply in the case.

In order to obtain an analytic solution, we choose the $1/c^2$ linear sound speed profile,

$$c_1(z) = c_1(1-az)^{-1/2}, \tag{34}$$

for a real constant a, with units of inverse length. The sound speed gradient at the interface is

$$g = \left.\frac{dc_1(z)}{dz}\right|_{z=0} = \frac{c_1 a}{2}. \tag{35}$$

Figure 13 shows the propagation paths for the gradient layer. In the environment studied in Fig. 14, where the sound speed at the top of the sediment is less than that in the water column, the refracted wave and the turning refracted wave are important contributors. Figure 14 shows the frequency dependence of the comparison between the far-field approximation and the numerical wave number integration for the case where there is a drop in sound speed at the water/sediment interface. The sound speed gradient g has a typical value 1.0 s^{-1}. The mismatch of the oscillation pattern is obvious at 100 Hz, but as the frequency increases, the match is

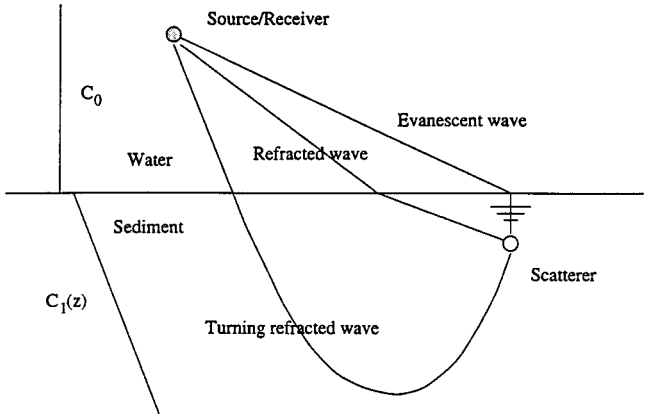


FIG. 13. The propagation paths for two half-space fast bottom with an upward-refracting sound speed profile in the bottom.

better. At 500 Hz, they are almost indistinguishable. The oscillation here comes from the interference between the direct incident wave and the upward-refracted wave. The sound speed gradient will determine the path length of the upward refracting ray. At high frequencies, the refracted wave will be attenuated so much that virtually no interference occurs. Therefore, the higher the frequency, the better the far-field approximation. The elucidation of the mismatch in oscillation patterns rests on the examination of the refracted wave. In the far-field approximation, the field is dominated by the interference between the plane wave incident at a particular angle and the refracted wave originating from the same angle of incidence. In the numerical wave number integration approach, the interference involves refracted waves originating from many incident angles.

The parameters in Fig. 15 correspond to the case where there is an increase in sound speed at the water/sediment interface. As a result, the evanescent wave comes into existence. The gradient g is again chosen to be 1.0 s^{-1}. Figure 15 shows the field for three different frequencies at two different depths. As we anticipated, the field now includes the contribution from the evanescent wave, the interface refracted wave, and the turning refracted wave, as shown in Fig. 13. Together, they will determine the interference pattern. The critical angle effect is still clear (cf. Sec. IV A). At higher frequencies, the results approach those for the two isovelocity half-space case because the turning refracted wave is strongly attenuated. In general, the far-field approximation will have some problems at low frequency for this kind of low-attenuation bottom with a sound speed gradient. For the high-attenuation sandy bottom we expect similar phenomena at low frequencies as those at high frequencies in Fig. 14. This behavior is demonstrated in Fig. 16. The parameters such as bottom sound speeds, density, and attenua-

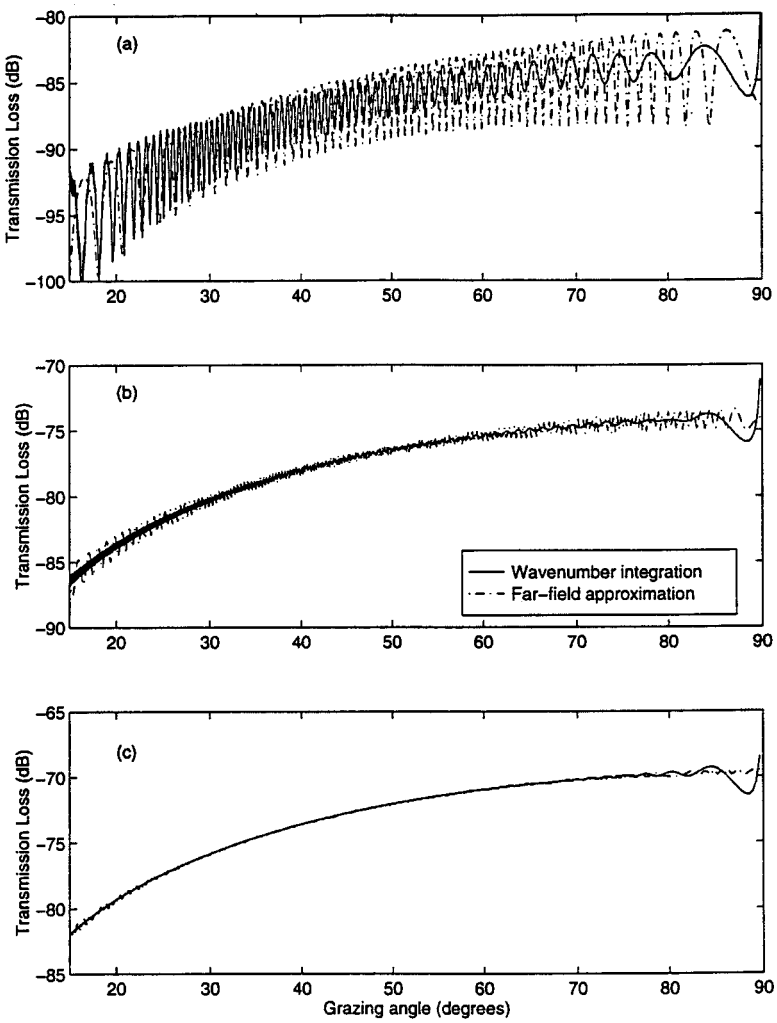


FIG. 14. The frequency dependence of the comparison results. Parameters are $c_0 = 1530$ m/s, $c_1 = 1510$ m/s, $\rho_1 = 1.72$ g/cm^3, $\delta = 0.001\,64$, $g = 1$ s^{-1}, $z_s = 100\lambda$, $z = 0$, and (a) $f = 100$ Hz, (b) $f = 300$ Hz, and (c) $f = 500$ Hz.

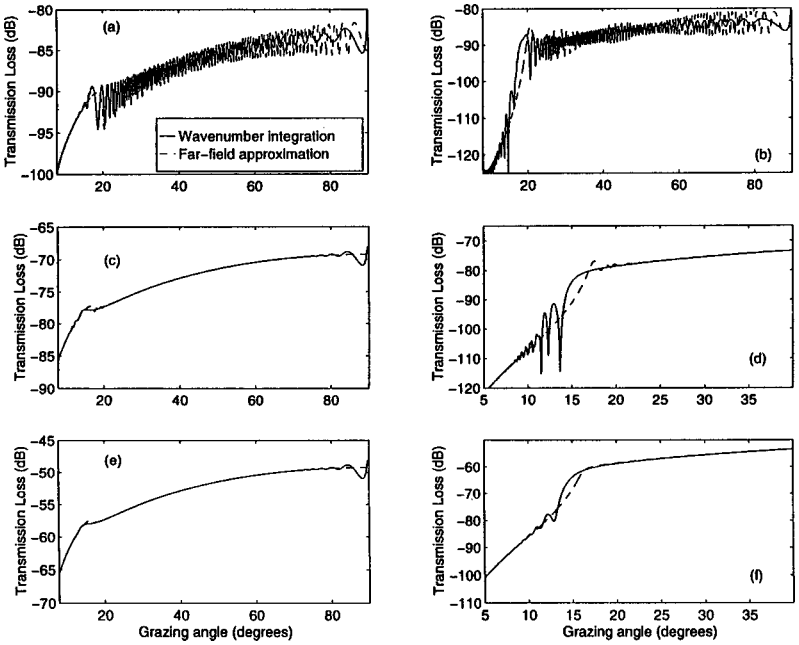


FIG. 15. The frequency and depth dependence of the comparison results for a low-attenuation bottom with a critical angle. Parameters are $c_0=1500$ m/s, $c_1=1560$ m/s, $\rho_1=1.8$ g/cm^3, $\delta=0.0027$, $g=1$ s^{-1}, $z_s=100\lambda$, and (a) $f=100$ Hz, $z=0$; (b) $f=100$ Hz, $z=-2\lambda$; (c) $f=500$ Hz, $z=0$; (d) $f=500$ Hz, $z=-2\lambda$; (e) $f=5000$ Hz, $z=0$; and (f) $f=5000$ Hz, $z=-2\lambda$.

tion coefficient in Fig. 16 are the same as for those in Fig. 10 except that the frequency is 100 Hz. Also, we have studied the effect of the sound speed gradient, which will determine the path length, and therefore the amplitude of the turning refracted wave in a bottom with attenuation. We would expect a higher amplitude for the turning refracted wave in a bottom with a larger gradient, because the path length would be shorter. As a result, the oscillation is more significant for a larger gradient as shown in Fig. 17.

The infinite sound speed profile will support multiple refracted paths that reflect off the seafloor interface and even a caustic at a specific range. However, the attenuation will minimize both effects.

Again we will study the influence of the propagation model on the quantification of the scattering process. We still assume point scatterers here and calculate only the backscattering strength. The scattering layer will be considered starting from the water/sediment interface only. For a high-attenuation sandy bottom like that used to generate Fig. 16, we choose σ_v to be 0.000 906. As shown in Fig. 18, the equivalent surface backscattering strength calculated with the far-field approximation is very close to the exact solution. For a low-attenuation, soft sediment bottom, we let the geoacoustic input be the same as that for Fig. 15(b) in Mourad and Jackson's paper.[11] Figure 19 shows the backscattering strength predicted by the far-field approximation and the numerical wave number integration technique. Notice that the oscillations are absent for the numerical solution but still present for the far-field approximation. An interesting phenomenon appearing in the results obtained by numerical wave number integration is that the backscattering strength has a drop of about 3 dB at 16 degrees, which is not shown in the results when the far-field approximation is applied. In Mourad and Jackson's Fig. 15(b) (cf. Fig. 20), the experimental data also show an 8–10 dB fall. Although the drop in our prediction does not match the data exactly, it is one step closer and suggests that the far-field approximation may not be good enough in situations like this.

C. Three layers with a sound speed gradient in the middle layer

A more realistic model of the sediment is to have a basement beneath the water/sediment interface. We still choose a $1/c^2$ linear sound speed profile below the water/sediment interface but constant sound velocity in the basement. Scatterers are assumed to be distributed in the middle layer.

The presence of an additional layer interface enables the insonification of scatterers through some extra paths, i.e., the single or multiple bounces between the two interfaces. Figure 21 shows the propagation paths between source/receiver and scatterer. For a slow bottom, there will be a shallow

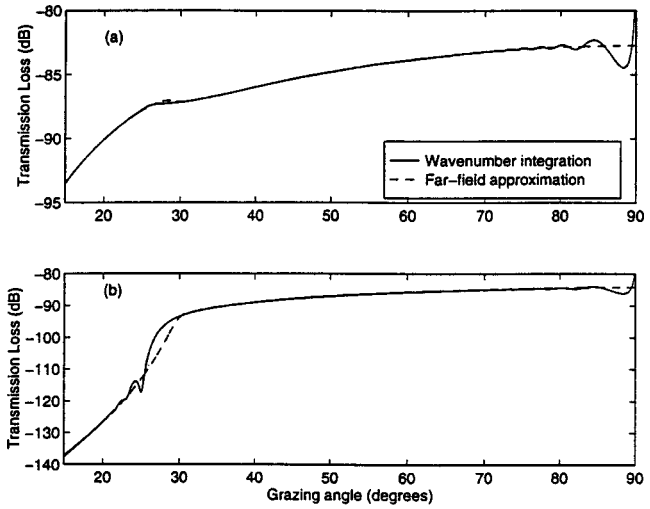


FIG. 16. The depth dependence for a high-attenuation bottom. Parameters are $f=100$ Hz, $c_0=1500$ m/s, $c_1=1689$ m/s, $\rho_1=1.97$ g/cm^3, $\delta=0.0166$, $g=1$ s^{-1}, and $z_s=100\lambda$. The transmitted fields at (a) $z=0$ and (b) $z=-2\lambda$.

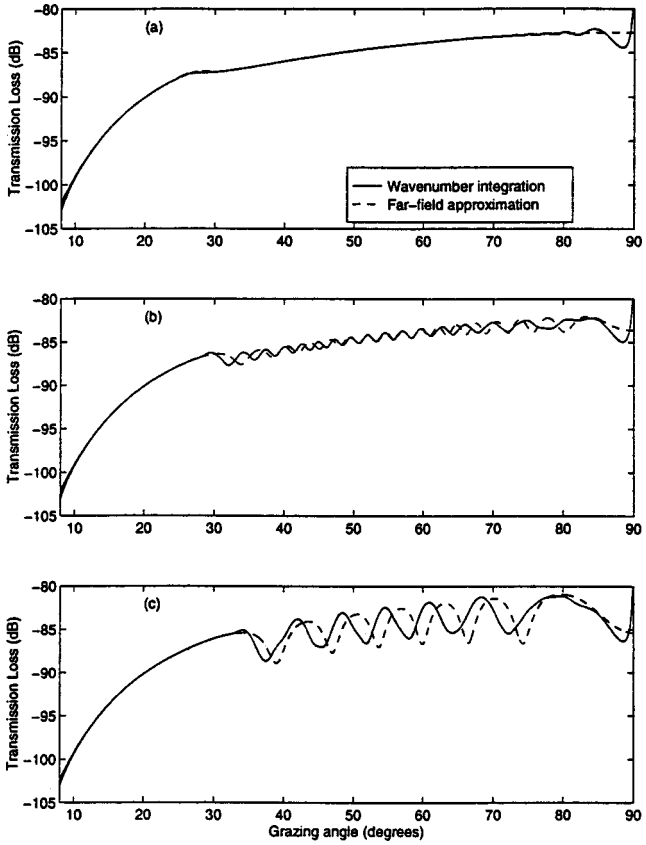


FIG. 17. The sound speed gradient dependence. Parameters are the same as those in Fig. 13(a) except the gradient (a) $g=1\ \mathrm{s}^{-1}$, (b) $g=4\ \mathrm{s}^{-1}$, and (c) $g=10\ \mathrm{s}^{-1}$.

angle refracted path and a steeper angle turning refracted or a reflected path from the subbottom layer, depending on the range (for given gradient and layer depth). For a fast bottom, there is an additional evanescent path for the subcritical ranges. For a very thick middle layer, the attenuation is likely to dampen all the multi-bounced refracted waves due to their long ray paths. For a thin middle layer, the turning

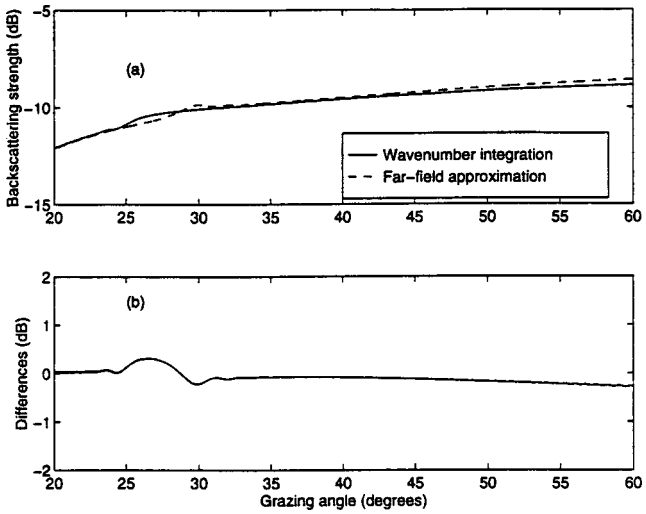


FIG. 18. The equivalent surface backscattering strength for a high-attenuation bottom assuming point scatterers. Parameters are the same as those in Fig. 16 with the starting depth of the scattering layer to be at $z=0$. (a) The backscattering strength and (b) differences.

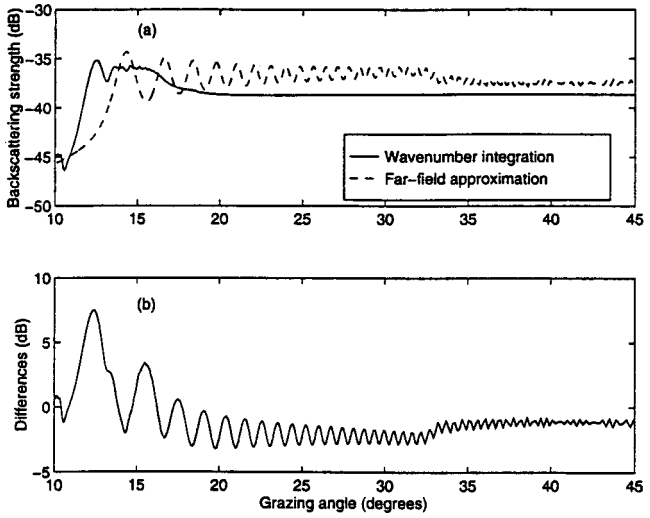


FIG. 19. The equivalent surface backscattering strength for a low-attenuation bottom assuming point scatterers. Parameters are $f=2300$ Hz, $c_0=1550$ m/s, $c_1=1576$ m/s, $\rho_1=1.83\ \mathrm{g/cm}^3$, $\delta=0.001\,93$, $g=10\ \mathrm{s}^{-1}$, and $z_s=100\lambda$. The scattering layer starts at $z=0$. (a) The backscattering strength and (b) differences.

refracted wave becomes the reflected wave off the subbottom interface. The attenuation is again a key factor in determining the contributions of multi-bounced refracted waves. The above scenario is similar to some degree to the situation where there is a sound speed gradient in the bottom. The turning point there can be considered to be the virtual lower boundary, although at a different depth for a different refracted angle.

As mentioned earlier, the middle layer thickness and attenuation are two key factors in determining the propagation field. One would gather that the middle layer thickness H is likely to play an important role in the interference between multipath rays. Figure 22 shows the field at the water/sediment interface for different middle layer thicknesses. The parameters for the middle layer are the same as those for Fig. 14(c) and correspond to a slow bottom. For the thin middle layer, the ''frequency'' of the oscillation is lower than that of the thick layer. However, the agreement between the far-field

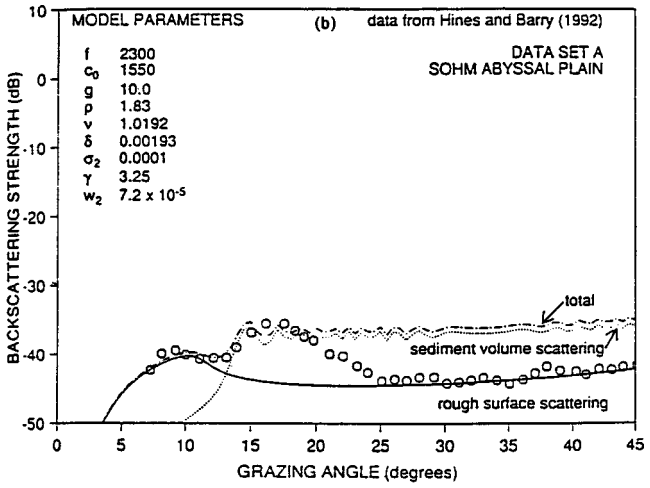


FIG. 20. Figure 15(b) in Mourad and Jackson's paper (Ref. 11). The circles here represent the data points.

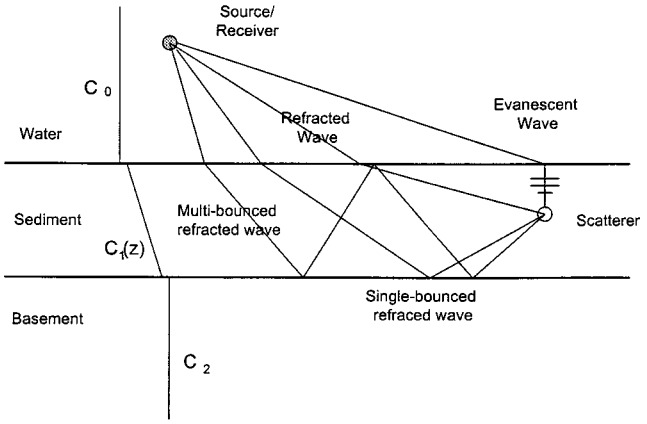


FIG. 21. Propagation paths for three layers with sound speed gradient in the middle layer.

approximation and the exact solution is better. The ''frequency'' of the oscillation is related to the rate of the phase change for the transmitted wave with the change of the incident angle. For a thicker middle layer, a small change of the incident angle would cause a big change of the phase of the transmitted then reflected wave, which results in high ''frequency'' oscillations. The interference pattern will depend on the phase of the reflected rays and therefore on the layer thickness.

Next we will see the effects of attenuation on the propagated field. The geoacoustic input for Fig. 23 for the upper two layers is the same as that for Fig. 12(b), corresponding to a sandy bottom. The attenuation coefficient is higher than in Fig. 22. Although the argument for the ''frequency'' of the oscillation still holds, the level of oscillation is much smaller because higher attenuation reduces the contributions of propagated waves with longer propagation paths. As can be seen in Fig. 23, the interference at a small grazing angle is hardly visible because the single- and multi-bounced refracted waves are too weak to interfere with the refracted

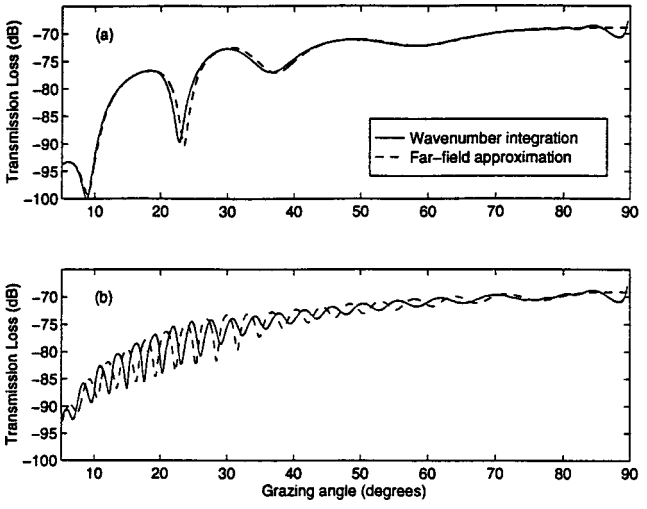


FIG. 22. The middle layer thickness dependence for a three-layer environment with a low-attenuation middle layer. Parameters are $f=500$ Hz, $c_0=1530$ m/s, $c_1=1510$ m/s, $c_2=1800$ m/s, $\rho_1=1.72$ g/cm^3, $\rho_2=1.9$ g/cm^3, $\delta_1=0.001\,64$, $\delta_2=0.019$, $g=1$ s^{-1}, $z_s=100\lambda$, $z=0$, and (a) $H=2\lambda$, and (b) $H=10\lambda$.

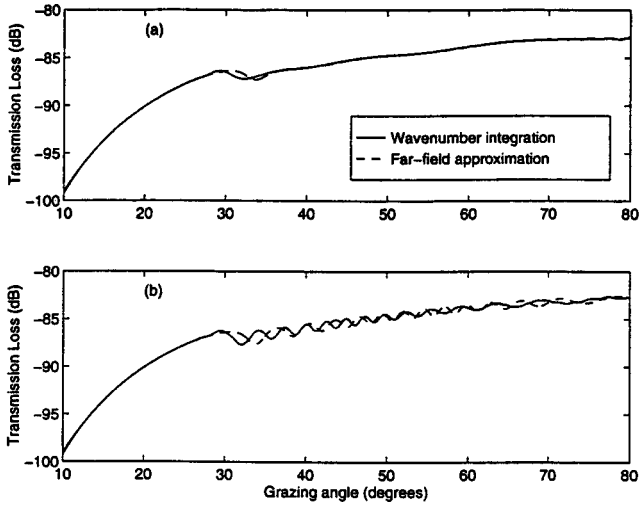


FIG. 23. The middle layer thickness dependence for a three-layer environment with a high-attenuation middle layer. Parameters are $f=100$ Hz, $c_0=1500$ m/s, $c_1=1689$ m/s, $c_2=1800$ m/s, $\rho_1=1.97$ g/cm^3, $\rho_2=2.2$ g/cm^3, $\delta_1=0.0166$, $\delta_2=0.019$, $g=4$ s^{-1}, $z_s=100\lambda$, $z=0$, and (a) $H=2\lambda$ and (b) $H=10\lambda$.

wave. As a result, the far-field approximation works better for a high-attenuation than for a low-attenuation bottom in the three-layer case.

V. CONCLUSIONS

In this article, we have compared several propagation models that are widely used in scattering studies with the exact solution obtained by a numerical wave number integration technique.

Hines' method, which decomposes the spherical wave into two plane waves, the refracted wave and the evanescent wave, matches the exact solution very well in a two-isovelocity half-space scenario. The overestimate of the field in a small region (less than 2 degrees) close to the critical angle is tolerable. Yet the difficulty in obtaining the stationary points easily and accurately hampers its generalization to more complex sound speed structure.

The far-field approximation, on the other hand, can be easily implemented in complicated bottom models, because it includes only one stationary point and has simplified its stationary point searching procedure by sacrificing some accuracy. As a result, the agreement with the exact solution is case dependent. In a two-isovelocity half-space model, the approach works almost perfectly if there is no critical angle at the water/sediment interface, which is likely to be the case for soft sediment types such as mud and silty clay. However, in some environments, the presence of the critical angle for such low-attenuation, soft sediment will yield some underestimate of the field at subcritical angles. The situation becomes worse for deeper scatterers because of the omission of the refracted wave at smaller than critical angle inherent in the far-field approximation. The high attenuation in sandy bottoms would ease the problem to some extent by attenuating the refracted wave due to its long path. As for the influence on the equivalent surface scattering strength calculation, results are encouraging if the scattering layer starts from the water/sediment interface. However, when the scattering

layer starts some distance away from the interface, which is unlikely for a sandy bottom but possible for soft sediments, the error introduced by applying the far-field approximation is significant.

The occurrence of the sound speed gradient in the sediment provides an additional propagation path, i.e., the turning refracted wave, which contributes to the determination of the interference pattern. The frequency-dependent attenuation and the gradient controls the amplitude and phase of the turning refracted wave. Therefore, the results of comparisons are frequency and gradient dependent. In general, the far-field approximation fares better for a high-attenuation, low-gradient bottom. For a low-attenuation bottom the difference is considerably larger, as can be seen in the equivalent surface scattering strength calculation.

In the three-layer case, multiple reflections between the two interfaces furthur complicate the multipath phenomena, causing oscillations of the field which are sensitive to the layer thickness no matter what type of sediment is in the middle layer. High bottom attenuation is again helpful in alleviating the multipath effect. The far-field approximation matches the numerical wave number integration better for a thinner middle layer and a higher attenuation situation.

The validity of using the equivalent surface scattering strength to characterize the volume scattering process is debatable at lower frequencies. Because of the evident multipath contribution to the field in many cases discussed in the comparison, the obtained angular dependence of the equivalent surface scattering strength is under question. However, as shown in our results, it is still reasonable to apply the equivalent surface scattering strength concept in some high-frequency and high-bottom-attenuation bottom scattering studies.

ACKNOWLEDGMENTS

The authors are grateful to Cynthia Sellers of WHOI for her help in many ways. This work was supported by the U.S. Office of Naval Research Contract No. N00014-96-1-0645. This article is Woods Hole Oceanographic Institution Contribution No. 9812.

[1] H. M. Merklinger, ''Bottom reverberation measured with explosive charges fired deep in the ocean,'' J. Acoust. Soc. Am. **44**, 508–513 (1968).
[2] D. R. Jackson, D. P. Winebrenner, and A. Ishimaru, ''Application of the composite roughness model to high-frequency bottom backscattering,'' J. Acoust. Soc. Am. **79**, 1410–1422 (1986).
[3] A. P. Lyons, A. L. Anderson, and F. S. Dwan, ''Acoustic scattering from the seafloor: Modeling and data comparison,'' J. Acoust. Soc. Am. **95**, 2441–2451 (1994).
[4] D. Tang, G. Jin, D. R. Jackson, and K. L. Williams, ''Analyses of high-frequency bottom and subbottom backscattering for two distinct shallow water environments,'' J. Acoust. Soc. Am. **96**, 2930–2936 (1994).
[5] D. Tang, G. V. Frisk, C. J. Sellers, and D. Li, ''Low-frequency acoustic backscattering by volumetric inhomogeneities in deep-ocean sediments,'' J. Acoust. Soc. Am. **98**, 508–516 (1995).
[6] R. J. Urick, *Principles of Underwater Sound* (McGraw–Hill, New York, 1983).
[7] D. R. Jackson, K. B. Briggs, K. L. Williams, and M. D. Richardson, ''Tests of models for high-frequency sea-floor backscatter,'' IEEE J. Ocean Eng. **21**, 458–470 (1996).
[8] P. A. Crowther, ''Some statistics of the sea-bed and acoustic scattering therefrom,'' in *Acoustics and the Sea-Bed*, edited by N. G. Pace (Bath Univ., Bath, England, 1983), pp. 147–155.
[9] T. Yamamoto, ''Acoustic scattering in the ocean from velocity and density fluctuations in the sediments,'' J. Acoust. Soc. Am. **99**, 866–879 (1996).
[10] A. N. Ivakin, ''Sound scattering by random inhomogeneities of stratified ocean sediments,'' Sov. Phys. Acoust. **32**, 492–496 (1986).
[11] P. D. Mourad and D. R. Jackson, ''A model/data comparison for low-frequency bottom backscatter,'' J. Acoust. Soc. Am. **94**, 344–358 (1993).
[12] P. C. Hines, ''Theoretical model of acoustic backscatter from a smooth seabed,'' J. Acoust. Soc. Am. **88**, 324–334 (1990).
[13] E. K. Westwood, ''Complex ray methods for acoustic interaction at a fluid–fluid interface,'' J. Acoust. Soc. Am. **85**, 1872–1884 (1989).
[14] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw–Hill, New York, 1953).
[15] D. Tang and G. V. Frisk, ''Spectral parametrization of sound scattering from a random ocean bottom,'' J. Acoust. Soc. Am. **92**, 2792–2799 (1992).
[16] A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, New York, 1978).
[17] G. V. Frisk, *Ocean and Seabed Acoustics: A Theory of Wave Propagation* (Prentice–Hall, Englewood Cliffs, NJ, 1994).
[18] H. Schmidt, SAFARI: Seismo-acoustic fast field algorithm for range independent environments, User's Guide. Rep. SR-113, 1988; SACLANT Undersea Research Centre, La Spezia, Italy.
[19] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, New York, 1993).
[20] E. L. Hamilton, ''Sound velocity and related properties of marine sediments, North Pacific,'' J. Geophys. Res. **75**, 4423–4446 (1970).
[21] A. C. Kibblewhite, ''Attenuation of sound in marine sediments: A review with emphasis on new low-frequency data,'' J. Acoust. Soc. Am. **86**, 716–738 (1989).
[22] R. D. Stoll, *Sediment Acoustics* (Springer-Verlag, New York, 1989).
[23] E. L. Hamilton, ''Compressional wave attenuation in marine sediments,'' Geophysics **37**, 620–646 (1972).
[24] A. K. Rogers, T. Yamamoto, and W. Carey, ''Experimental investigation of sediment effect on acoustic wave propagation in the shallow ocean,'' J. Acoust. Soc. Am. **93**, 1747–1761 (1993).
[25] E. L. Hamilton, ''Sound velocity gradients in marine sediments,'' J. Acoust. Soc. Am. **65**, 909–922 (1979).

A comparison of bistatic scattering from two geologically distinct abyssal hills

Chin Swee Chia[a] and Nicholas C. Makris[b]
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Laurie T. Fialkowski
Naval Research Laboratory, Washington, D.C. 20375

(Received 22 September; revised 14 June 2000; accepted 16 June 2000)

The bistatic scattering characteristics of two geologically distinct abyssal hills located on the western flank of the Mid-Atlantic Ridge, known as B′ and C′, are experimentally compared using data acquired with low-frequency towed-array systems at $\frac{1}{2}$ convergence zone (~33 km) stand-off. The comparison is significant because the abyssal hills span the two classes of elevated seafloor crust that cover the Mid-Atlantic Ridge. The highly lineated B′ feature is representative of abyssal hills composed of outside corner crust, the most commonly occurring category, whereas the domed C′ promontory is representative of the rougher, low-aspect-ratio abyssal hills composed of inside corner crust. The latter are less common and usually restricted to segment valley margins. The mean biazimuthal scattering distributions of the two abyssal hills each exhibit Lambertian behavior with comparable albedos, suggesting that the distinction between abyssal hills composed of differing crust is not significant in modeling long-range reverberation. The adverse effect of using bathymetry that undersamples seafloor projected area in scattering strength analysis is also quantified with data from the B′ ridge. Specifically, the use of undersampled bathymetry can lead to significant overestimates in the strength of seafloor scattering. © *2000 Acoustical Society of America.* [S0001-4966(00)04410-6]

PACS numbers: 43.30.Hw, 43.30.Gv [DLB]

I. INTRODUCTION

In this paper, a comparison is made of *bistatic* scattering from two geologically distinct abyssal hills located along a segment valley on the western flank of the Mid-Atlantic Ridge (MAR). This analysis is based on acoustic data acquired by low-frequency towed-array systems at $\frac{1}{2}$ convergence zone (CZ) (~33 km) stand-off during the Main Acoustics Experiment (MAE) of the Acoustic Reverberation Special Research Program (ARSRP) in July 1993.[1–4] A detailed analysis of bistatic scattering from one of these abyssal hills, named B′, has previously been presented in Ref. 3. The goal of this paper is to present a similar analysis for the second ridge, C′, for comparative purposes. The comparison is significant because the B′ and C′ abyssal hills span the two classes of elevated seafloor crust that cover the MAR. An acoustic analysis of these two prominent bathymetric features should then lead to a better understanding of the long-range, bistatic scattering properties of bathymetric highs throughout the MAR.

The B′ abyssal hill is composed of *outside corner* (OC) crust, whereas the C′ promontory is composed of *inside corner* (IC) crust.[5] This geological distinction has already helped to clarify measured differences in the spatial characteristics of *monostatic* reverberation from these two features.[4] The lineated B′ feature, by its high aspect ratio, is representative of abyssal hills composed of OC crust. The numerous steep escarpments that run along the major axis of OC abyssal hills, and B′ in particular, return echoes that faithfully image the lineated scarp morphology.[3,4] The domed C′ promontory is representative of the rougher, low-aspect-ratio abyssal hills composed of IC crust. Long-range acoustic images of C′ again faithfully image steep slopes on C′, but these show more amorphous, nonlinear structure.[4] Abyssal hills of OC crust occur more commonly, while those of IC crust are usually restricted to segment valley margins. Geologically, ''IC crust forms on the side of the spreading axis next to an active discontinuity and is characterized by anomalously shallow bathymetry, thinned crust and/or mantle exposures, irregular large-throw normal faults, and a paucity of volcanic morphological features. OC crust is formed on the opposite side of the spreading axis next to the inactive trace of the discontinuity; it has more normal depth and crustal thickness, regular fault patterns, and more common volcanic features.''[5]

In a previous analysis of ARSRP data,[3] high-resolution bistatic reverberation images of B′ were generated from $\frac{1}{2}$ CZ stand-off. These measured images were compared with the modeled images, generated from 5-m resolution bathymetry data, to show that steep scarps on B′ return the strongest echoes because they project the largest surface areas along the acoustic path from source to scattering patch to receiver. Both measured and modeled images also show that prominent echo returns deterministically image the scarp morphology when the cross-range resolution of the towed-array system runs along the scarp axis. Although small-scale features along the scarp, such as canyons and gullies (~100–200-m scale), are theoretically resolvable in range by the towed-

[a]Current address: DSO National Laboratories, 20 Science Park Drive, S118230, Singapore.
[b]Electronic mail: makris@mit.edu

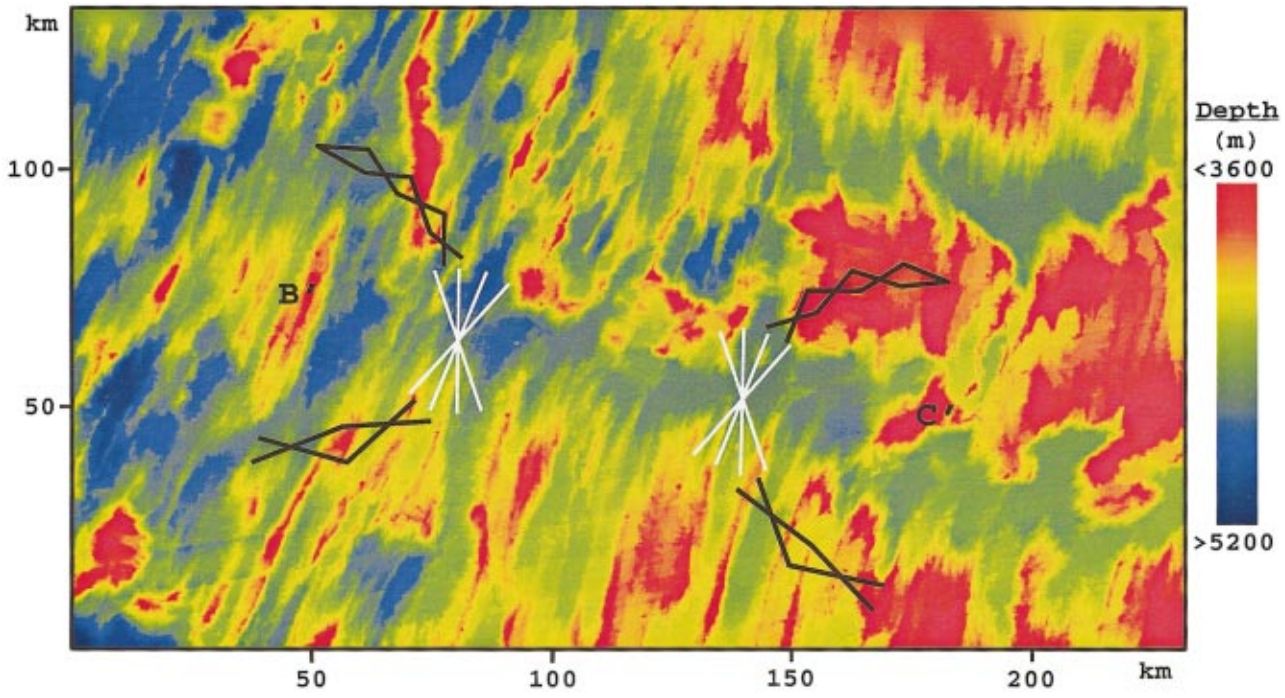


FIG. 1. Bistatic tow-ship tracks overlain on the 200-m resolution bathymetry data of the experimental area, extracted from Ref. 4. The B′ and C′ ridges are two prominent seafloor features at opposite ends of the segment valley that runs roughly east–west across the experimental area. White tracks denote the monostatic positions of RV CORY CHOUEST that trace the Easternstar and Westernstar, while the black tracks indicate the bistatic positions of RV ALLIANCE along the semicircular arcs about B′ and C′.

array system at some bistatic angles, statistical fluctuations due to signal-dependent noise present in the actual data prevent the system from resolving these features. This leads to the conclusion that signal-dependent noise, known as speckle, is one of the primary factors limiting the towed-array system's resolving power in imaging the seafloor geomorphology.[3,4,6]

In the same study, the biazimuthal scattering distribution functions[3] of the two major scarps on B′ were estimated using 5-m resolution bathymetry data. The mean strengths of the biazimuthal scattering distributions over the two scarps were shown to be identical and equal to the constant of -17 dB $\pm$ 8 dB. This lead to the hypothesis that long-range reverberation from prominent geomorphologic features of the world's Mid-Ocean Ridges may be adequately modeled as Lambertian with albedo $\pi/10$[1.7].

To further test this hypothesis, a similar study has been carried out to measure the biazimuthal scattering distribution function of a major scarp on the C′ abyssal hill. Since 5-m resolution bathymetry is unavailable at C′, the analysis is carried out with lower-resolution hydrosweep bathymetry data sampled at 200-m intervals.[7] To help control the comparison with C′ results, and quantify the potentially adverse effects of using undersampled bathymetry data in scattering strength estimation, the biazimuthal scattering distribution function for the B′ scarps are recomputed using the lower-resolution hydrosweep data. While a number of investigators have analyzed monostatic reverberation from B′[3,4,8,9] and bistatic reverberation from B′,[3,4] this is the first study[10] to analyze bistatic reverberation from C′.

II. BISTATIC EXPERIMENTAL DESIGN AND GEOMORPHOLOGY OF THE TWO OCEAN RIDGES

The experiments took place within a subsection of the ONR Natural Laboratory spanning 25.5° to 27.5° North latitude and 45° to 49° West longitude along the western flank of the MAR.[1–5] Nine experiments, referred to as *runs* in the ARSRP community, were conducted. Bistatic scattering was measured at $\frac{1}{2}$ CZ stand-off from site B′ in runs 5a and 5b and from C′ in runs 3 and 8. The B′ and C′ abyssal hills are separated by a segment valley of roughly 2 CZ length which enabled a set of experiments at $\frac{1}{2}$ and 1 $\frac{1}{2}$ CZ to be conducted about each feature with extreme efficiency.[2,4,11] These experiments are therefore referred to as the B′–C′ *corridor* experiments, which comprised roughly 90% of the Main Acoustics Experiment.[1–4,11]

The experiments were conducted using two research vessels (RVs), the CORY CHOUEST and ALLIANCE.[12] Their bistatic tow-ship tracks are overlain on the local bathymetry in Fig. 1. During each run, CORY and ALLIANCE began at the edges of the star-shaped tracks with slow cruising speeds of 3.0–4.5 knots. While the CORY traced its straight-line path in the central star, the RV ALLIANCE zigzagged along semicircular arcs about B′ and C′. To maximize the sonar cross-range resolution at B′ and C′, the towed-array's broadside beam was directed towards each target abyssal hill, while the ships' radiated noise was restricted to the lowest-resolution endfire beams to minimize mutual noise interference.

It has been shown[3,4,13] that significant variations in reverberation can occur for small changes in measurement position due to bathymetry-induced variations in transmission loss (TL). The star-shaped ship tracks avoid the problem of

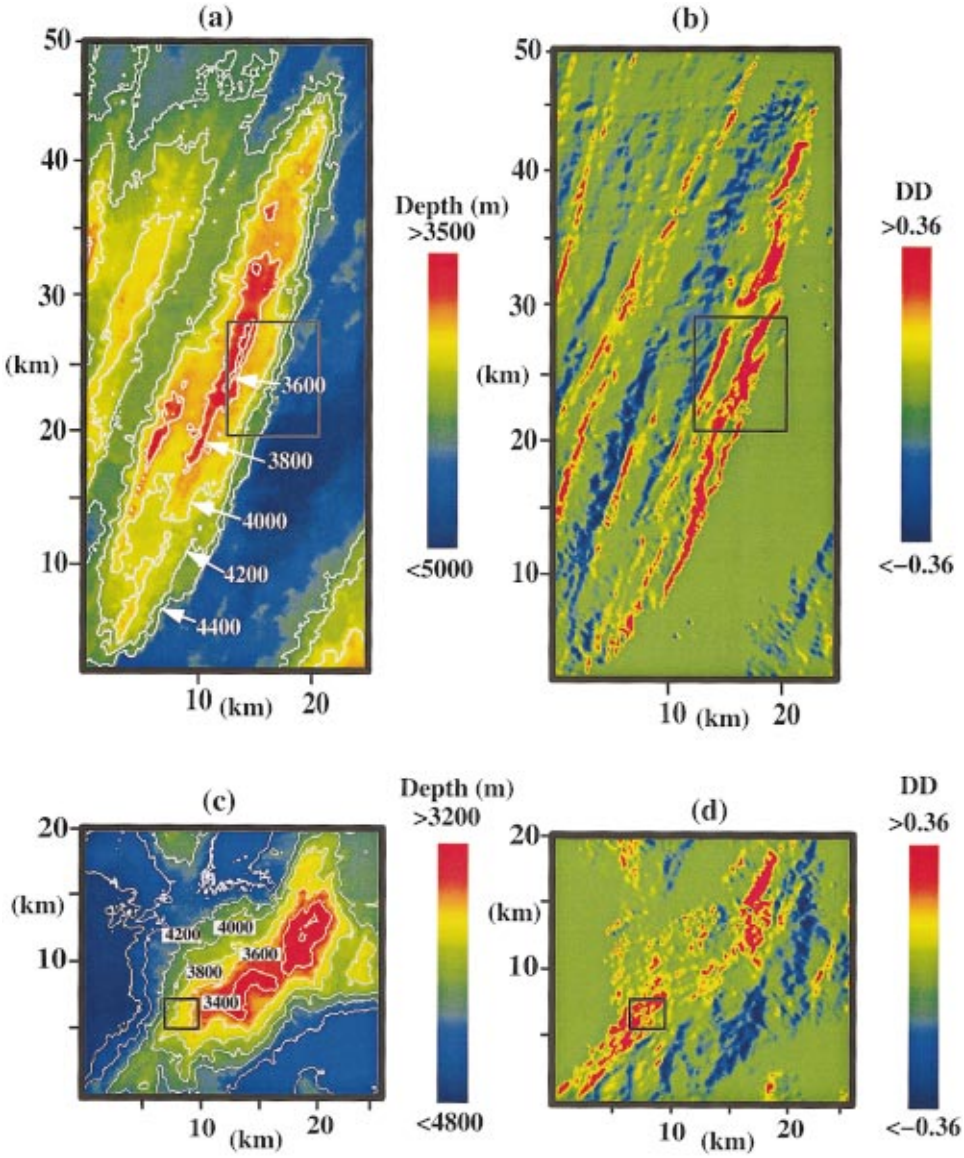


FIG. 2. Comparison between of the B′ and C′ bathymetric features and their directional derivatives (DD) charts. (a) The bathymetry of the B′ abyssal hill plotted at 200-m resolution. (b) The DD of B′ with respect to source at Easternstar center. (c) The bathymetry of the C′ abyssal hill plotted at 200-m resolution. (d) The DD of C′ with respect to source at the Easternstar center.

comparing measurements with different TL by providing a point of global convergence over all towed-array headings at the star centers. The CORY tracks for runs 3 and 8, at $\frac{1}{2}$ CZ from C′, and 5a and 5b, at $\frac{1}{2}$ CZ from B′, are therefore referred to as the *Easternstar* and *Westernstar*, respectively.[4]

The CORY transmitted from a ten-element vertical line array (VLA) with on-axis source level calibrated to roughly 229 dB *re* 1 μPa@1 m. This source was deployed with center at 181 m. It transmitted a variety of waveforms, including the linear frequency modulation (LFM) waveform, which swept across the 200–255-Hz frequency band in 5 s, that is used exclusively in the present analysis. During each LFM transmission interval, acquired reverberation data were assigned a *data segment number* to identify the corresponding transmission cycle. Reverberation returned from each ocean ridge was received by the CORY's 128-element horizontal line array (HLA) at 170-m depth, in an effectively monostatic manner since the CORY source and receiver arrays were separated by roughly 1.12 km from array center to array center. The ALLIANCE HLA was towed at an average depth of 460 m for the bistatic receptions. Although the ALLIANCE source was also deployed in these experiments, we did not analyze data associated with it since its source strength and directivity was much lower than that of CORY. Specifically, only the CORY LFM transmissions, at the center frequency of 227.5 Hz, have been analyzed since they provide the best range resolution of ~14 m. The match-filtered data are averaged over 0.0625 s for CORY's receptions and 0.0533 s for ALLIANCE's receptions, and the effective range resolutions for the two towed-array systems are computed to be roughly 47 and 40 m, respectively.[4]

The geomorphology and gradient components of the B′ and C′ abyssal hills are shown in Fig. 2. While B′ has the classic elliptical shape, with high aspect ratio and long lineated scarps running parallel to its ridge axis, that typifies most abyssal hills of OC crust, C′ appears dome-like and dominated by normal faults with variable orientations, as is common among abyssal hills of IC crust. The structural difference between these two abyssal hills of distinct class can be better illustrated by bathymetric slope gradients components along the path sound travels or directional derivatives (DDs).[3,4,13] The DD is here defined as the inner product of

the bathymetric gradient with a local unit vector pointing in the horizontal direction of the source or receiver. In Fig. 2(b), where the DD is taken with respect to the Easternstar Center, the two major scarps on the eastern face of B′ appear prominently along the abyssal hill's major axis as lineations of high positive DD. By contrast, as shown in Fig. 2(d) where DD is again taken roughly with respect to the Easternstar Center, positive DD of C′ appears speckled and irregularly scattered over the entire ridge in accordance with its irregularly oriented faults. Besides the structural differences, the crustal composition of B′ and C′ is expected to be different since inside corners like C′ typically consist of plutonic rocks and mantle ultramafics such as peridotites and serpentinites,[5] while B′ is comprised of basalt thickly coated with iron–manganese.[7]

III. THE EFFECT OF BATHYMETRIC UNDERSAMPLING ON SCATTERING DISTRIBUTION ESTIMATION AT B′

A. Comparison of the 5-m resolution and 200-m resolution images

The effect of bathymetric undersampling on bistatic scattering analysis is evaluated using two different resolutions of bathymetry data for the B′ abyssal hill. Lower resolution (200-m sampled) hydrosweep data is taken from the precise 8×9 km region on the east-central face of B′ where the high-resolution (5-m sampled) bathymetry data are available. The exact 5-m resolution contours, for DD$>\frac{1}{2}$, used in Ref. 3 to designate the B′ scarps, were mapped point-to-point onto the 200-m grid chart. For comparison, the shaded relief plots at 5-m and 200-m resolutions for a section of the upper scarp are shown in Figs. 3(a) and (b), respectively. In the 5-m resolution plot, steep slopes and small 200-m scale anomalies, such as canyons and gullies, are clearly observed along the upper scarp. In the 200-m resolution plot, the upper scarp appears relatively flat over the sonar resolution footprint,[3] and small-scale anomalies are not properly resolved. Typical slopes on the B′ upper and lower scarps exceed 50° according to the 5-m sampled bathymetry, as shown in Fig. 2(b) of Ref. 3 and Fig. 22 of Ref. 4, but are significantly underestimated in the 200-m resolution bathymetric data set where the maximum slopes are found to be roughly 20°.

Two-way transmission loss (TL) and surface projection factors are computed over the resultant 8×9-km region using the same method as described in Ref. 3 but now with the lower resolution bathymetry sampled at 200 m. The parabolic equation is used to compute the two-way TL and a ray-trace method is used to model refraction due to depth-dependent sound-speed variations in the water column so that two-way travel time can be converted to range for reverberation charting. Rays are also traced to determine the incident angle θ_i from the source to the seafloor patch and the scattered angle θ_r from the seafloor to the receiver. These angles are measured relative to the seafloor normal. The seafloor's surface projection terms in the direction of incident and scattered rays are computed as in Ref. 3 via $C_i = 10\log(\cos\theta_i)$ and $C_r = 10\log(\cos\theta_r)$, respectively. To provide an illustrative example, surface projection and two-way TL charts obtained using bathymetry data sampled at 5-m and 200-m resolutions are shown in Fig. 4 for data segment S435.

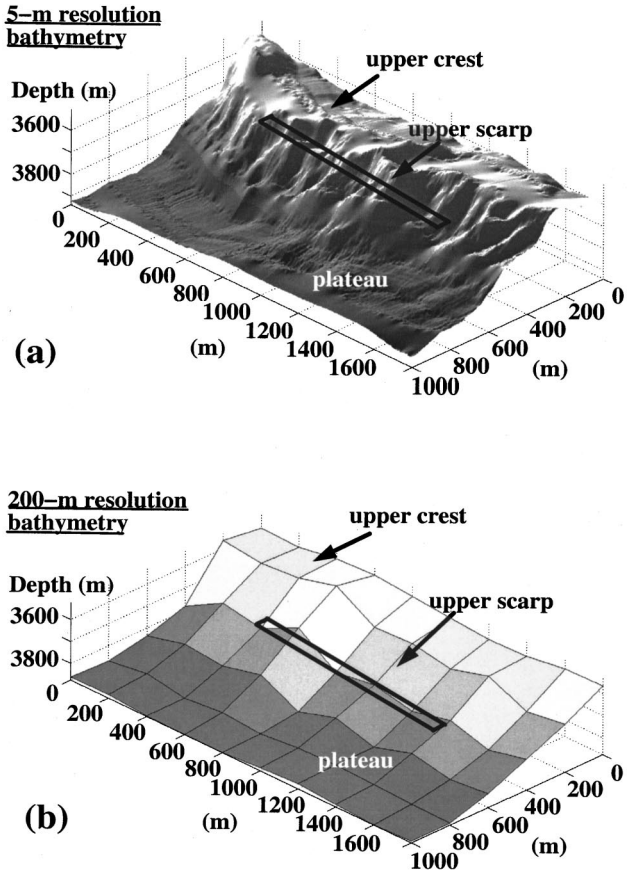


FIG. 3. Shaded relief plots for a section of the upper scarp at (a) 5-m resolution, and (b) 200-m resolution. Overlain is a typical sonar resolution footprint for monostatic reception at $\frac{1}{2}$ CZ with receiving array parallel to ridge axis. While the small-scale anomalies along the scarps such as canyons and gullies are clearly observed at 5-m resolution, these anomalies are not properly resolved at 200-m resolution. Slopes on such scarps are found to be severely underestimated in the 200-m resolution bathymetry.

Figures 4(a) and (b) illustrate the surface projection, $C_i = 10\log(\cos\theta_i)$, computed at two resolutions over the designated 8×9-km scarp area. Since the upper and lower scarps of B′ have slopes that typically exceed 50°, their surface projections in the 5-m resolution chart appear mostly in red, which correspond to values close to 0 dB. The surface projections of the two scarps in the 200-m resolution chart are well below 0 dB due to their underestimated slopes. The plateaus register with extremely low surface projections, as expected. Figures 4(c) and (d) compare two-way TL charted for S435 monostatic reception at 200-m and 5-m resolutions, respectively. While the details are lost when lower-resolution bathymetry is used, and these same details are averaged over in wide-area towed-array resolution footprints, the overall TL levels are similar for both high- and low-resolution bathymetry.

B. Measured and modeled reverberation at B′

Modeled reverberation is computed by the same method and for the same regions and segments as in Ref. 3, again

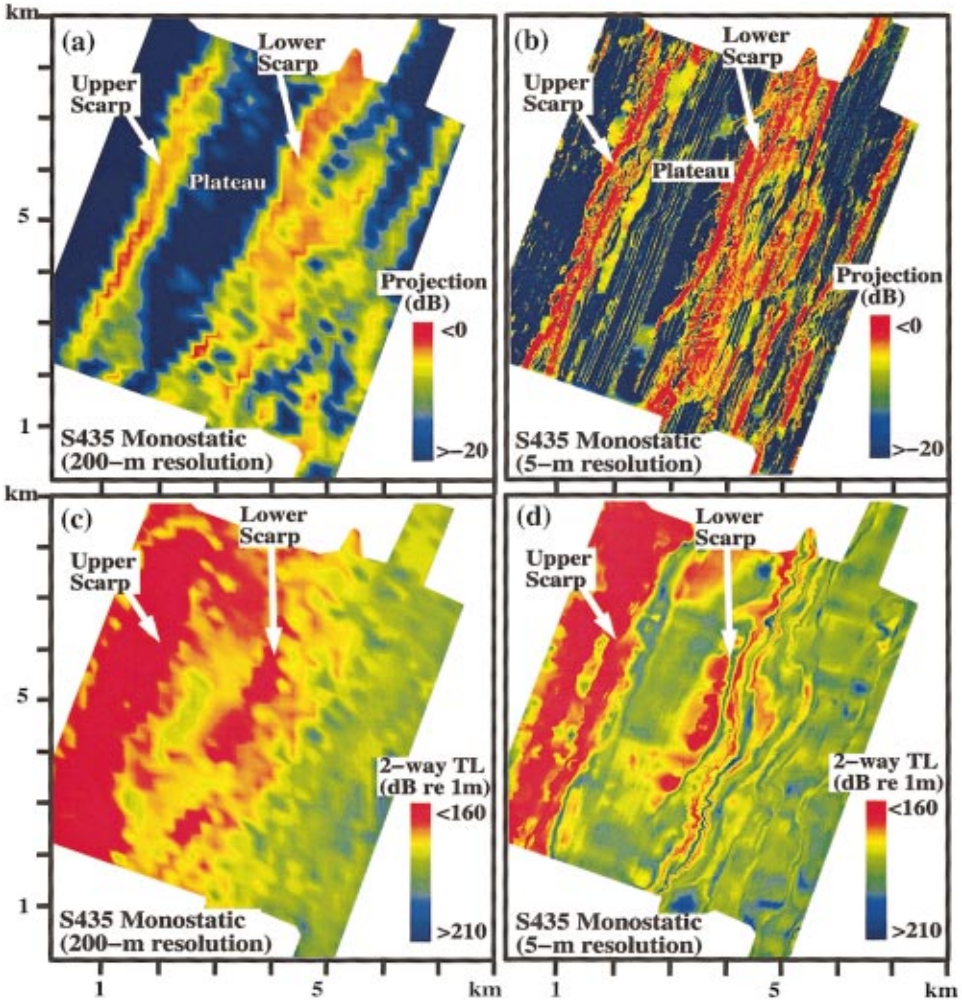


FIG. 4. Surface projection, $C_i = 10\log(\cos\theta_i)$, computed over the upper and lower scarps of B′ using (a) 200-m resolution, and (b) 5-m resolution bathymetry data for S435. The two steep scarps of B′ are undersampled at 200-m resolution and so, besides appearing blurry, yield significantly lower surface projections than at 5-m resolution. (c) and (d) show the two-way TL charts computed over the same site for 200-m and 5-m resolutions, respectively, for S435. Both the transmission losses computed over the scarp area are found to be similar in magnitudes, although detail is lost at 200-m resolution.

assuming a perfectly reflecting Lambertian surface, with unity albedo, except that 200-m sampled bathymetry is used instead of the 5-m sampled bathymetry. Since the 200-m grid size exceeds the towed-array range resolution of 40–47 m, there is no averaging of data over bathymetric range cells when performing the spatial convolution at 200-m resolution. Consequently, the modeled reverberation over the two scarps, shown for S435 in Fig. 5(b), has a more speckled appearance in the 200-m resolution charts than in the corresponding 5-m resolution charts, as shown for S435 in Fig. 12 of Ref. 3.

Prominent measured and modeled returns from the scarp areas show a good correlation at 200-m resolution, as shown for example in Fig. 5, just as they do for 5-m resolution, as shown for the same segment S435 in Fig. 12 of Ref. 3. While the model predicts strong lineated echoes to be returned along the scarp axes in the monostatic reception, Fig. 5(b), a speckle-like echo pattern across the two scarps is predicted in the corresponding bistatic reception, in Fig. 5(d), as described previously in Ref. 3. The general character of prominent returns measured over the two scarps, as illustrated in Figs. 5(a) and (c), agrees well with the predictions.

C. Biazimuthal scattering distributions of the two B′ scarps

As demonstrated in Ref. 3, the scarp elevation within the sonar resolution footprint cannot be approximated as a planar surface since there is no unique surface normal to characterize the multiple bathymetric features within the resolution footprint. It is therefore meaningless to plot the estimates of scattering strength as a function of incident and scattered angles. Instead, the statistical approach adopted in Ref. 3 is used to describe the mean scattering distribution over the B′ scarps as a function of receiver azimuth Ω_r, with respect to the normal horizontally bisecting the B′ scarp axis. The strength of the biazimuthal scattering distribution is averaged over the designated area according to Eq. (A5) of the Appendix. A full biazimuthal description of the scattering distribution over the B′ scarps, with respect to source and receiver azimuths, can be regained by referring to the distribution plot of source–receiver location pairs in Fig. 3 of Ref. 3. The biazimuthal scattering distribution strength is computed just as in Ref. 3 except that 200-m resolution bathymetry data is used instead of 5-m resolution data.

In Figs. 6(a) and (b), the curves of the mean measured reverberation level, at 200-m resolution, over the upper scarp $\langle R(x,y|\Omega_i,\Omega_r)\rangle_{A_{\rm up}}$ and the lower scarp $\langle R(x,y|\Omega_i,\Omega_r)\rangle_{A_{\rm low}}$ are plotted as a function of receiver azimuth Ω_r together with their standard deviations $\sigma_{A_{\rm up}}\{R(x,y)\}$ and $\sigma_{A_{\rm low}}\{R(x,y)\}$. The measured reverberation curve is computed via Eq. (A2) in the Appendix, and the subscripts $A_{\rm up}$ and $A_{\rm low}$ denote that the reverberation levels are averaged over the upper scarp and lower scarp areas, respectively. For

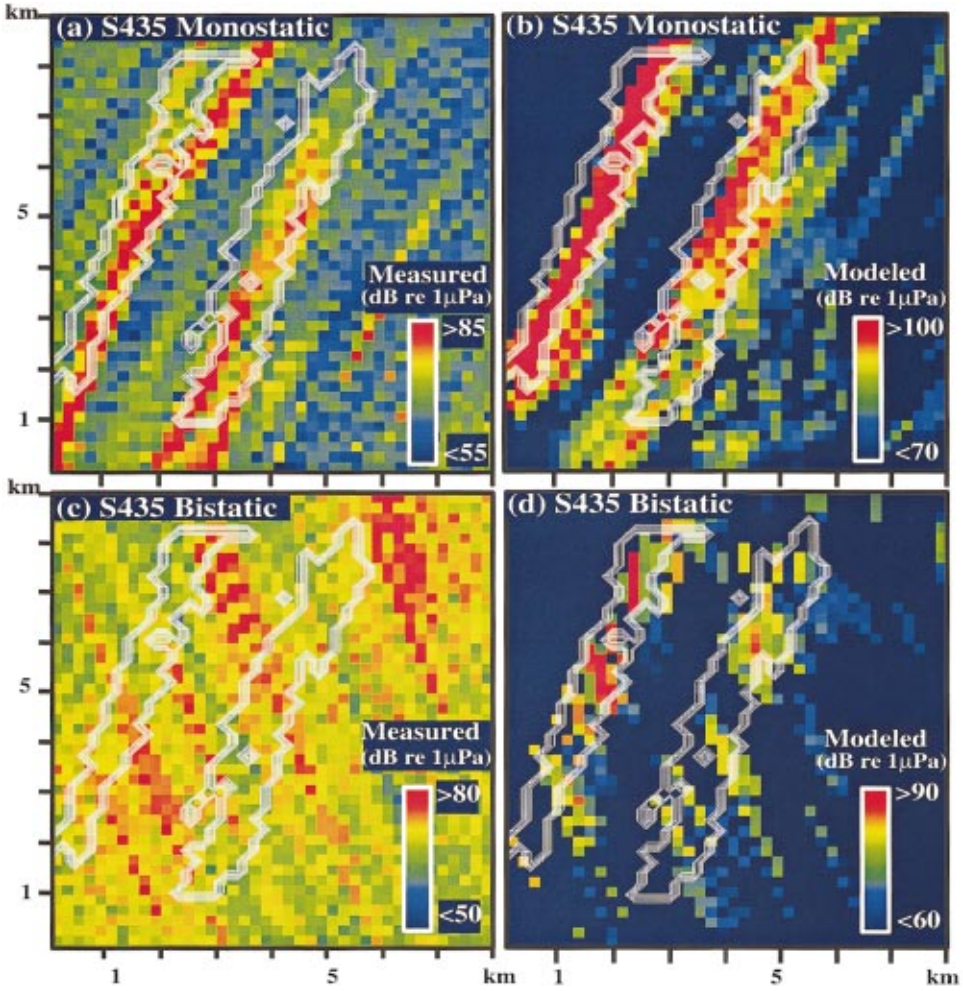


FIG. 5. Charts of monostatic and bistatic measured and modeled reverberation for S435 over the upper and lower scarp contours at 200-m resolution. (a) Measured monostatic reverberation. (b) Model monostatic reverberation. (c) Measured bistatic reverberation. (d) Modeled bistatic reverberation.

comparisons, the mean values of the 5-m resolution curves, previously obtained in Ref. 3, are plotted in dotted lines overlain in the same figure. Generally, these reverberation curves at 5-m and 200-m resolutions are found to be relatively constant across the $\pm 90°$ receiver azimuths, and their respective mean values closely match.

Curves of the mean reverberation level modeled over the two scarps, $\langle R_M(x,y|\Omega_i,\Omega_r)\rangle_{A_{\text{up}}}$ and $\langle R_M(x,y|\Omega_i,\Omega_r)\rangle_{A_{\text{low}}}$, are plotted as a function of receiver azimuth Ω_r in Figs. 7(a) and (b). The modeled reverberation curve is obtained using Eq. (A3) of the Appendix. Again, the modeled reverberation curves at 200-m resolution (in solid lines) exhibit the same trend as the 5-m resolution curves (in dotted lines) across receiver azimuth but are uniformly lower by several dB. While the mean values for the upper scarp at both resolutions fluctuate within $|\Omega_r|<30°$, the mean values for the lower scarp display a slight convex behavior with peak values within $|\Omega_r|<30°$ and roll off by $\sim$10 dB towards the extreme azimuths. It can also be easily seen that the 200-m resolution curves for the two scarps are distinctly lower than their corresponding 5-m resolution curves. The offset between the 200-m and 5-m resolution curves can be explained by examining the surface projection and two-way TL prior to the spatial convolution.

Figure 8 shows the mean surface projections $\langle C_i(x,y|\Omega_i,\Omega_r)+C_r(x,y|\Omega_i,\Omega_r)\rangle_{A_{\text{up}}}$ and $\langle C_i(x,y|\Omega_i,\Omega_r)+C_r(x,y|\Omega_i,\Omega_r)\rangle_{A_{\text{low}}}$ over the upper and lower scarps as a function of receiver azimuth Ω_r. The surface projection curves, at 200-m resolution, exhibit the same convex behavior as the 5-m resolution curves, with their peak values near to the origin. However, the 200-m resolution curves are found to be roughly 6 to 8 dB lower than the mean surface projections at 5-m resolution. This offset is directly caused by the use of the undersampled 200-m resolution bathymetry data, where the projected area of the B′ scarps, with respect to the refracted ray paths, is highly underestimated.

The mean two-way TL, $\langle \text{TL}_i(x,y|\Omega_i,\Omega_r)+\text{TL}_r(x,y|\Omega_i,\Omega_r)\rangle_{A_{\text{up}}}$ and $\langle \text{TL}_i(x,y|\Omega_i,\Omega_r)+\text{TL}_r(x,y|\Omega_i,\Omega_r)\rangle_{A_{\text{low}}}$, are plotted as a function of receiver azimuth Ω_r in Fig. 9. The two-way TL curves, at 200-m resolution, match the 5-m resolution curves almost precisely, both in terms of their mean values and standard deviations at each receiver azimuth. The difference observed in the modeled reverberation curves at 5-m and 200-m sampling is then a direct consequence of the surface projection underestimate caused by the use of undersampled bathymetry.

The mean strengths of the biazimuthal scattering distributions estimated at 200-m resolution over the upper scarp $\langle \hat{\tilde{F}}(x,y|\Omega_i,\Omega_r)\rangle_{A_{\text{up}}}$ and lower scarp $\langle \hat{\tilde{F}}(x,y|\Omega_i,\Omega_r)\rangle_{A_{\text{low}}}$ are plotted as a function of receiver azimuth Ω_r in Fig. 10, along with their standard deviations $\sigma_{A_{\text{up}}}\{\hat{\tilde{F}}(x,y)\}$ and

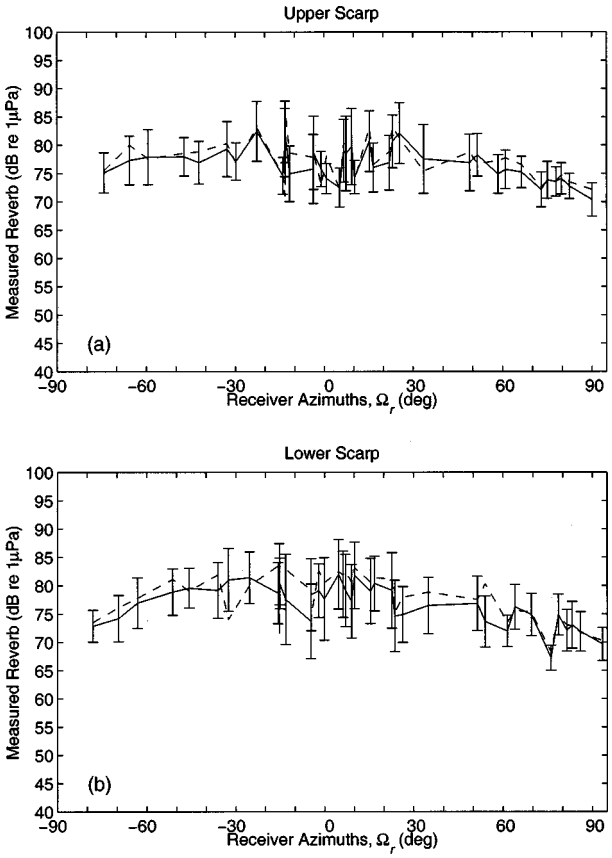


FIG. 6. Mean reverberation levels measured over (a) the upper scarp $\langle R(x,y|\Omega_i,\Omega_r)\rangle_{A_{\rm up}}$ and (b) the lower scarp $\langle R(x,y|\Omega_i,\Omega_r)\rangle_{A_{\rm low}}$ as a function of receiver azimuth Ω_r along with their respective standard deviations $\sigma_{A_{\rm up}}\{R(x,y)\}$ and $\sigma_{A_{\rm low}}\{R(x,y)\}$. Solid line denotes the 200-m resolution curve, and dotted line denotes the 5-m resolution curve.

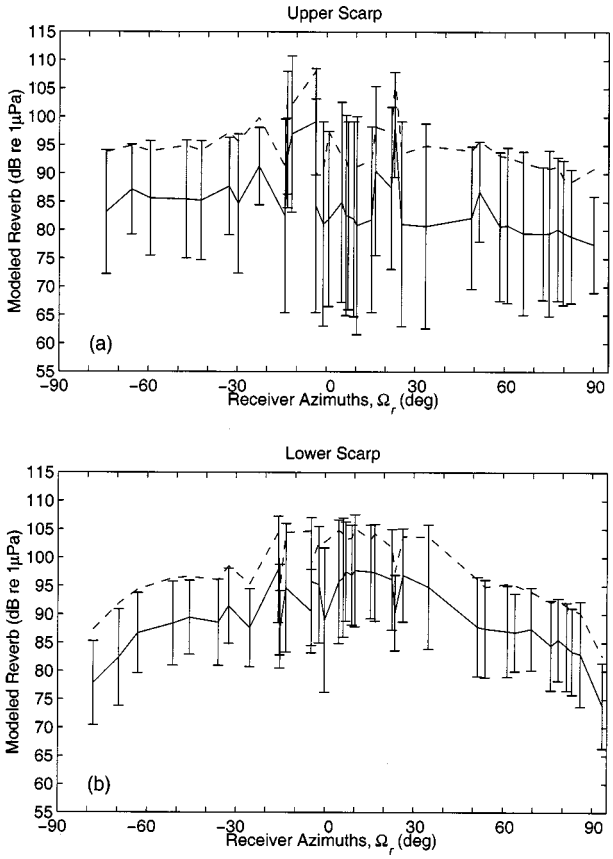


FIG. 7. Mean reverberation levels modeled over (a) the upper scarp $\langle R_M(x,y|\Omega_i,\Omega_r)\rangle_{A_{\rm up}}$, and (b) the lower scarp $\langle R_M(x,y|\Omega_i,\Omega_r)\rangle_{A_{\rm low}}$ as a function of receiver azimuth Ω_r along with their respective standard deviations $\sigma_{A_{\rm up}}\{R_M(x,y)\}$ and $\sigma_{A_{\rm low}}\{R_M(x,y)\}$. Solid line denotes the 200-m resolution curve, and dotted line denotes the 5-m resolution curve.

$\sigma_{A_{\rm low}}\{\hat{\bar{F}}(x,y)\}$. The 200-m resolution curves again show the same constant trend as the 5-m resolution curves but are uniformly about 6 dB higher in level. Specifically, a constant line can be drawn within the error bars across receiver azimuth for curves of both upper and lower scarp scattering distribution strength for both 200-m resolution and 5-m resolution results, but the line is at roughly -11 dB in the former and -17 dB in the latter case. This overestimate in scattering distribution strength at 200-m resolution is a direct consequence of the underestimate in surface projected area caused by use of the undersampled 200-m resolution bathymetry. The use of bathymetry that undersamples the projected area of the seafloor within the resolution footprint of the towed-array system can lead to significant overestimates in the strength of seafloor scattering.

IV. ANALYSIS OF BISTATIC SCATTERING FROM C′ WITH 200-m RESOLUTION BATHYMETRY

A. Experiment geometry

The 3×2-km region designated for this study of bistatic scattering from the C′ abyssal hill is overlain on 200-m sampled bathymetry in Fig. 11(a). This site at the southwest corner of the roughly 10×20-km C′ is selected because it is consistently insonified by the main beam of the CORY source array throughout the experiments. As shown in the magnified plots of Fig. 11(a), the selected 3×2-km C′ region is part of a steep scarp that faces northwest with depths ranging from 4000 to 3500 m that include the source's conjugate depth of 3800 m[3,4,13] and so intersects the refractive path of sound from sources and to receivers at $\frac{1}{2}$ CZ stand-off range. The directional derivative of C′, shown in Fig. 11(b), is computed with the source located to the west of C′ within 1 km of the Eastern Star center at the CORY's position during S229 transmissions. The DD exhibits a speckled pattern consistent with the irregularly oriented faults known to characterize the geomorphology of C′.

The locations of the two research vessels for data segments analyzed in the C′ study are plotted in Fig. 12. The central black box in this figure indicates the 3×2-km area at the southwest (SW) corner of C′ designated for this study, to be referred to as the SW box. A total of 10 monostatic and 22 bistatic segments has been analyzed to cover a full suite of 180° bistatic angles distributed in a semicircle at roughly $\frac{1}{2}$ CZ radius from the center of SW box. The boxed alphabets indicate the RV CORY's locations, while the corresponding unboxed alphabets denote the RV ALLIANCE's locations for the same transmission.

The biazimuthal distribution of the source–receiver lo-

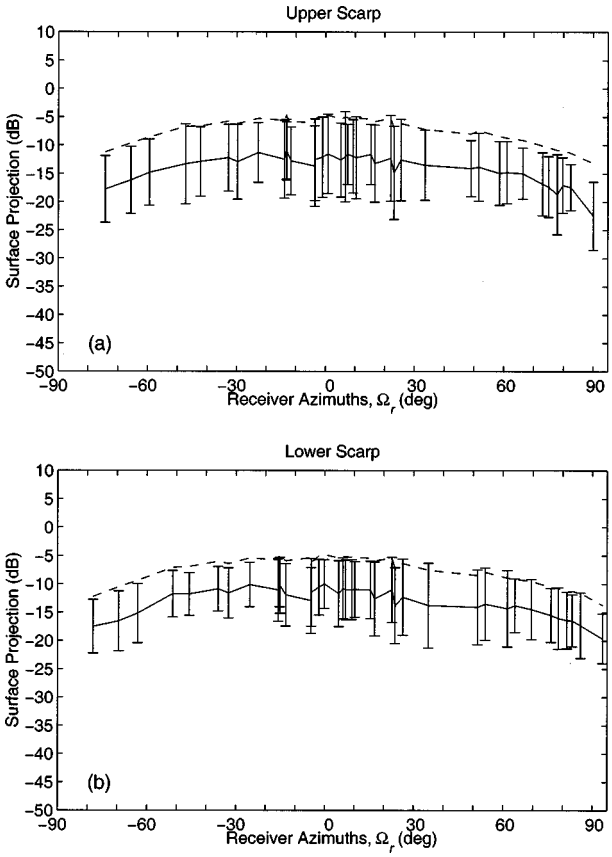


FIG. 8. The mean surface projection over (a) the upper scarp $\langle C_i(x,y|\Omega_i,\Omega_r)+C_r(x,y|\Omega_i,\Omega_r)\rangle_{A_{\text{up}}}$, and (b) the lower scarp $\langle C_i(x,y|\Omega_i,\Omega_r)+C_r(x,y|\Omega_i,\Omega_r)\rangle_{A_{\text{low}}}$ as a function of receiver azimuth Ω_r along with their standard deviations. Solid line denotes the 200-m resolution curve, and dotted line denotes the 5-m resolution curve.

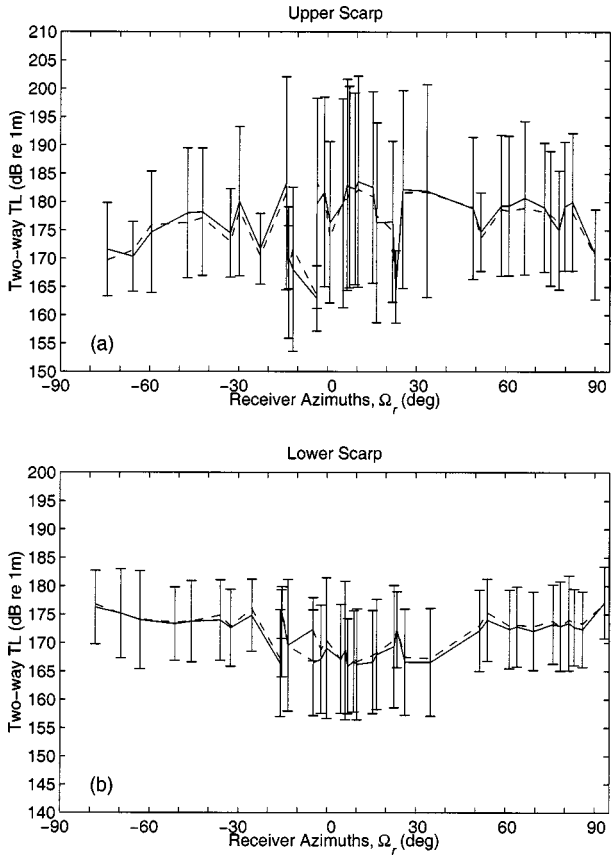


FIG. 9. The mean two-way transmission loss over (a) the upper scarp $\langle \text{TL}_i(x,y|\Omega_i,\Omega_r)+\text{TL}_r(x,y|\Omega_i,\Omega_r)\rangle_{A_{\text{up}}}$, and (b) the lower scarp $\langle \text{TL}_i(x,y|\Omega_i,\Omega_r)+\text{TL}_r(x,y|\Omega_i,\Omega_r)\rangle_{A_{\text{low}}}$ as a function of receiver azimuth Ω_r along with their standard deviations. Solid line denotes the 200-m resolution curve, and dotted line denotes the 5-m resolution curve.

cation pairs is plotted in Fig. 13. Azimuth is measured counterclockwise from a northwest line that originates at the center of the SW box and is normal to the scarp axis. Specifically, the source azimuth $\Omega_i=0°$ falls on this line which connects the RV CORY's location for segment S229 to the center of the SW box, as indicated by the dotted line between F and C′ in Fig. 12 which bisects the *Easternstar* ship tracks. While the source azimuths Ω_i fall within $|\Omega_i|<30°$, the receiver azimuths Ω_r span over a $\pm 90°$ sector for a complete study of biazimuthal scattering at C′.

B. Wide-area bistatic images

Monostatic and bistatic reverberation charts for S229, S874, and S883 are shown in Figs. 14, 15, and 16, respectively. ALLIANCE was located west of C′ in S874, midway to the northern extreme of its course in S229, and near the southern extreme in S883, with source–receiver pairs distributed according to $\Omega_i\sim 0°$, $\Omega_r\sim -57°$ for S229, $\Omega_i\sim 25°$, $\Omega_r\sim 37°$ for S874, and $\Omega_i\sim 7°$, $\Omega_r\sim 69°$ for S883. These distinct bistatic locations along with their distinct towed-array headings lead to reverberation charts that are highly representative of the various geometrical issues at play in the present experiment.

In Figs. 14(a) to 16(a), prominent echoes are primarily charted over the western scarps of C′ in the vicinity of the conjugate depth contour. Since the CORY's source and its towed-array receiver are close to each other, ~1.2 km, with respect to their respective $\frac{1}{2}$ CZ ranges to C′, reception by CORY is effectively monostatic and so reverberation charts exhibit circular symmetry about the source/receiver location. Echo returns are ambiguously mirrored across the CORY receiver's axis due to the inherent left–right ambiguity of the linear towed array.[3,4,13,14] Differences in charted reverberation across the segments arise primarily from changes in TL and projected area associated with changes in CORY position.

The bistatic reverberation charts, in Figs. 14(c), 15(c), and 16(c), illustrate three typical scenarios in bistatic sonar reception. In S874, a circularly symmetric pattern arises as in monostatic reception since the separation between the RV CORY and ALLIANCE is relatively short (~6 km) compared to the range to C′. Moreover, the ambiguous returns are reflected across the ALLIANCE receiving array axis with near-perfect symmetry. In S883, reverberation arriving at the same travel time follows elliptical arcs about foci located at the well-separated source and receiver locations. Left–right ambiguity is relatively symmetric for S883 because the ALLIANCE receiver's heading coincides with the source–receiver axis. The ambiguous image of C′ then occupies a similar spatial area as the true one. In S229, the separation

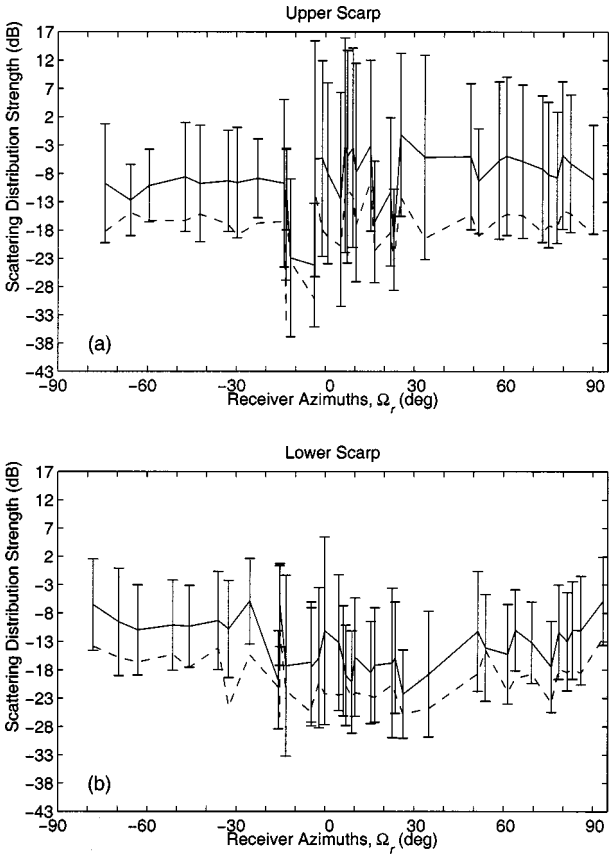


FIG. 10. The mean strength of the biazimuthal scattering distribution estimated over (a) the upper scarp $\langle \hat{\hat{F}}(x,y|\Omega_i,\Omega_r)\rangle_{A_{\rm up}}$, and (b) the lower scarp $\langle \hat{\hat{F}}(x,y|\Omega_i,\Omega_r)\rangle_{A_{\rm low}}$ as a function of receiver azimuth Ω_r along with their standard deviations $\sigma_{A_{\rm up}}\{\hat{\hat{F}}(x,y)\}$ and $\sigma_{A_{\rm low}}\{\hat{\hat{F}}(x,y)\}$. Solid line denotes the 200-m resolution curve, and dotted line denotes the 5-m resolution curve.

between CORY and ALLIANCE is again significant, but the ALLIANCE heading departs considerably from the source–receiver axis. This leads to an absence of symmetry about the receiver axis. The distortion compresses the C′ ambiguity to a much smaller spatial region than the true return. The C′ ambiguity also falls at a shorter range than the true return to preserve the two-way travel time. Similar behavior in wide-area bistatic reverberation charts has been documented previously at the B′ abyssal hill.[3]

Figures 14 to 16(b) and (d) illustrate the bistatic horizontal projection of bathymetry (BHBP), as defined in Ref. 3, computed over C′. Overlying the BHBP images are high-amplitude reverberation contours for the specified segment. Most of the western scarps of C′ register positive BHBP in the monostatic charts due to the CORY's predominantly western location. Prominent echoes register well with regions of positive BHBP. The BHBP for the bistatic charts varies significantly over the chosen three segments because of their differing source–receiver orientations. In the extreme case of S883, Fig. 16(d), only the SW corner of C′ yields positive BHBP, and consequently prominent returns, because the ALLIANCE has moved to the southwest of C′ and other portions of the abyssal hill are shadowed. The SW box is in this region and is almost always well insonified by the CORY's source and at the same time is acoustically visible to the ALLIANCE receiving array throughout the bistatic experiment at C′.

Wide-area images for all the data segments analyzed in this study have been examined to ensure that the inherent left–right ambiguity of the linear towed array did not corrupt the measured results over the designated SW box of C′. Although a few data segments, such as S220, S919, and

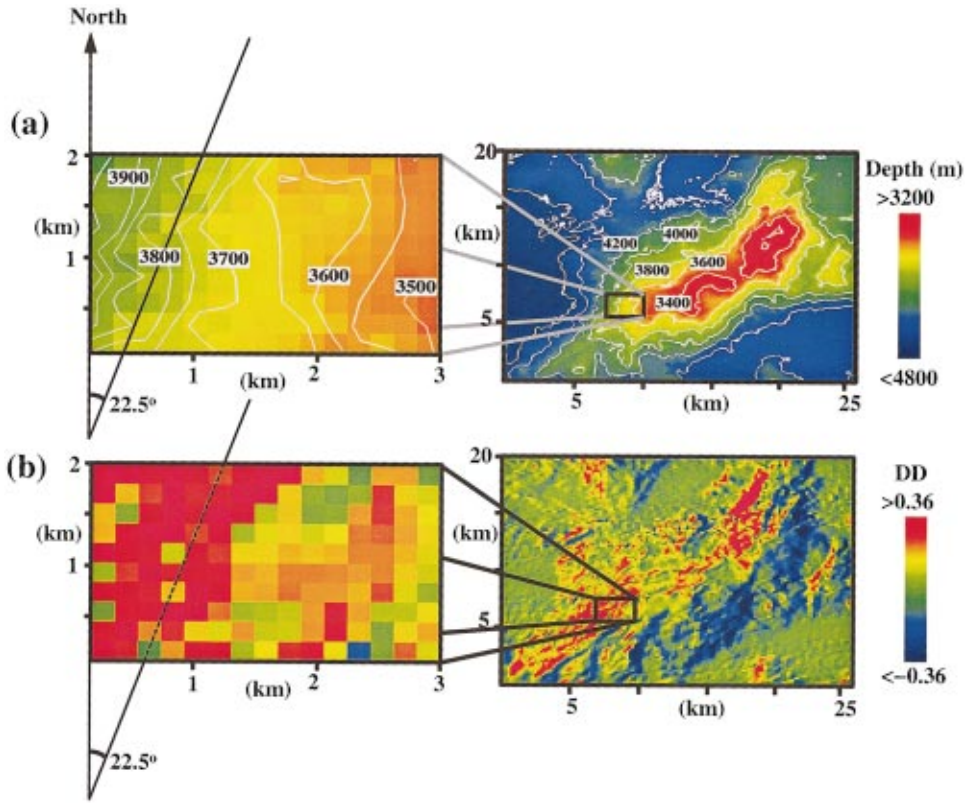


FIG. 11. (a) Bathymetry of the C′ inside corner abyssal hill sampled at 200-m intervals. The black box at the SW corner of C′ indicates the 3 ×2-km region designated for this study and subsequently referred to as the SW box. (b) The DD for C′ with respect to a source at the Easternstar center. Steep scarps facing the source are charted in red with DD>0.36, equivalent to slope gradients >20°; steep slopes facing away from the source are in blue. The axis of the scarps is at roughly 22.5° as shown.

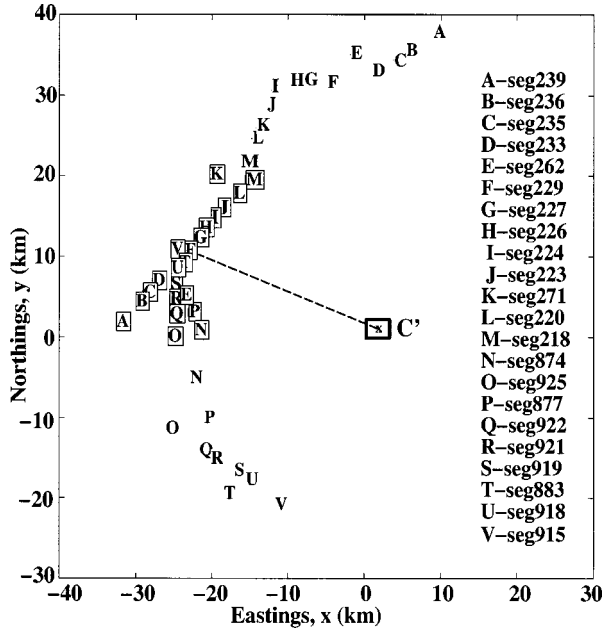


FIG. 12. Bistatic locations of the two research vessels during the LFM transmissions analyzed in the C′ study, given in Eastings and Northings. These locations are distributed in a semicircle about the center of SW box, which is the scarp area designated for the present study. The boxed alphabets denote the CORY's locations, i.e., the source locations, while the corresponding unboxed alphabets along the circular arc denote the ALLIANCE's locations, i.e., bistatic receiver locations, for the given transmission cycle.

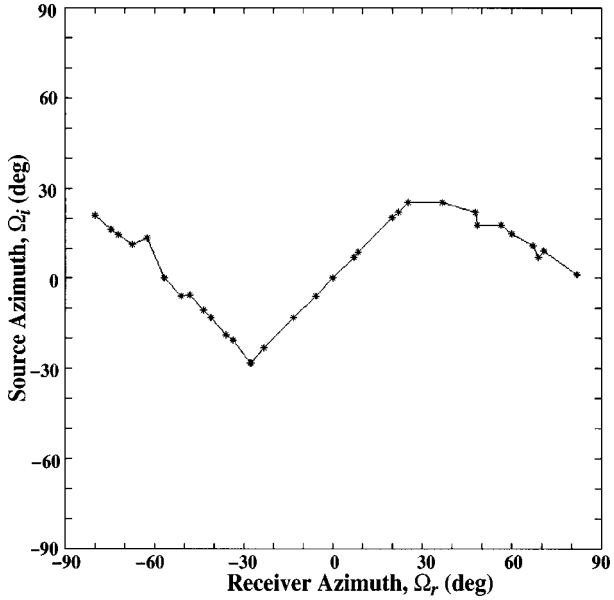


FIG. 13. Azimuthal distribution of the source and receiver pairs for the C′ study. While the source azimuths Ω_i fall within $|\Omega_i|<30°$, the receiver azimuths Ω_r span a full 180° range of nonforward azimuths about C′.

S925 monostatic segments, were found to have some ambiguities charted to the C′ ridge, these did not fall within the SW box.

C. Measured and modeled reverberation at C′

Figure 17 shows the surface projection and TL at the SW corner of C′ for segment S229. Since the designated 3 $\times$2-km SW box cannot adequately display broad spatial variations of TL due to bathymetry, a 9$\times$6-km area is illustrated here to provide more perspective. Specifically, the SW box shown as the central white box in all the figures indicates the scarp area at C′ designated for the present analysis. Transmission loss from source to scattering patch, TL_i, and scattering patch to receiver, TL_r, is produced by sweeping the broadband TL maps, which are incoherently averaged over 200–255-Hz band as in Ref. 3, across the bathymetry of C′ at 200-m resolution. The surface projection terms from source to seafloor, C_i, and seafloor to receiver, C_r, are obtained by sweeping across the scarp area with grazing angle maps produced by ray trace, as in Ref. 3.

Figures 17(a) and (b) illustrate the surface projections within the SW box along ray paths directed towards the source and receiver's locations. In segment S229, the RV CORY is located at the SW corner's broadside, while the RV ALLIANCE is midway to the northern extreme of its path. As a result, the SW corner scarp of C′ projects larger surface area towards the CORY than towards the ALLIANCE. The areas of extremely low surface projection, in dark blue in the lower right corners of Figs. 17(a) and (b), are in the shadow zone of the refracted sound paths.

The transmission loss charts TL_i and TL_r in Figs. 17(c) and (d) illustrate the distinct natures of broadband TL maps for source versus receiver. Figure 17(c) illustrates well-structured main beam behavior. The SW box has a low TL_i since it falls within the source's main beam, while the higher elevation above it suffers a high TL due to the shadow zone of the source main beam's refractive path. Figure 17(d) shows TL_r, from seafloor to the receiver, to be relatively constant across the site. Spatial variations in TL_i are found to be more dominant than TL_r in dictating the characteristics of the two-way TL across the SW corner of C′.

The measured and modeled reverberation charts for segments S229, S874, and S883 are presented in Figs. 18–20, respectively. Generally, prominent measured returns from the SW corner scarp of C′ show reasonably good agreement with corresponding modeled returns in the monostatic reverberation charts. Across these three segments, the character of reverberation changes predictably as a function of spatial variations of the TL and surface projection. For example, the model predicts correctly that the lower elevation of the SW corner scarp at C′ will return prominent echoes in S229 and S883 monostatic receptions, while it predicts higher elevations to return prominent echoes in S874 monostatic reception.

However, such a good visual correlation is sometimes not found in the bistatic reverberation charts over the SW corner box. Frequently, the modeled reverberation predicts strong echo returns from a specific area of the scarp, while the measured reverberation appears more diffusely scattered over the entire scarp area. This inconsistency has been explained in our B′ high-resolution study.[3] Specifically, the signal-dependent speckle noise arising from statistical fluctuations of the scattered field is sufficient to obscure the expected echo patterns. That is, within the insonified scarp, the detailed structure of bistatic returns has variations on the

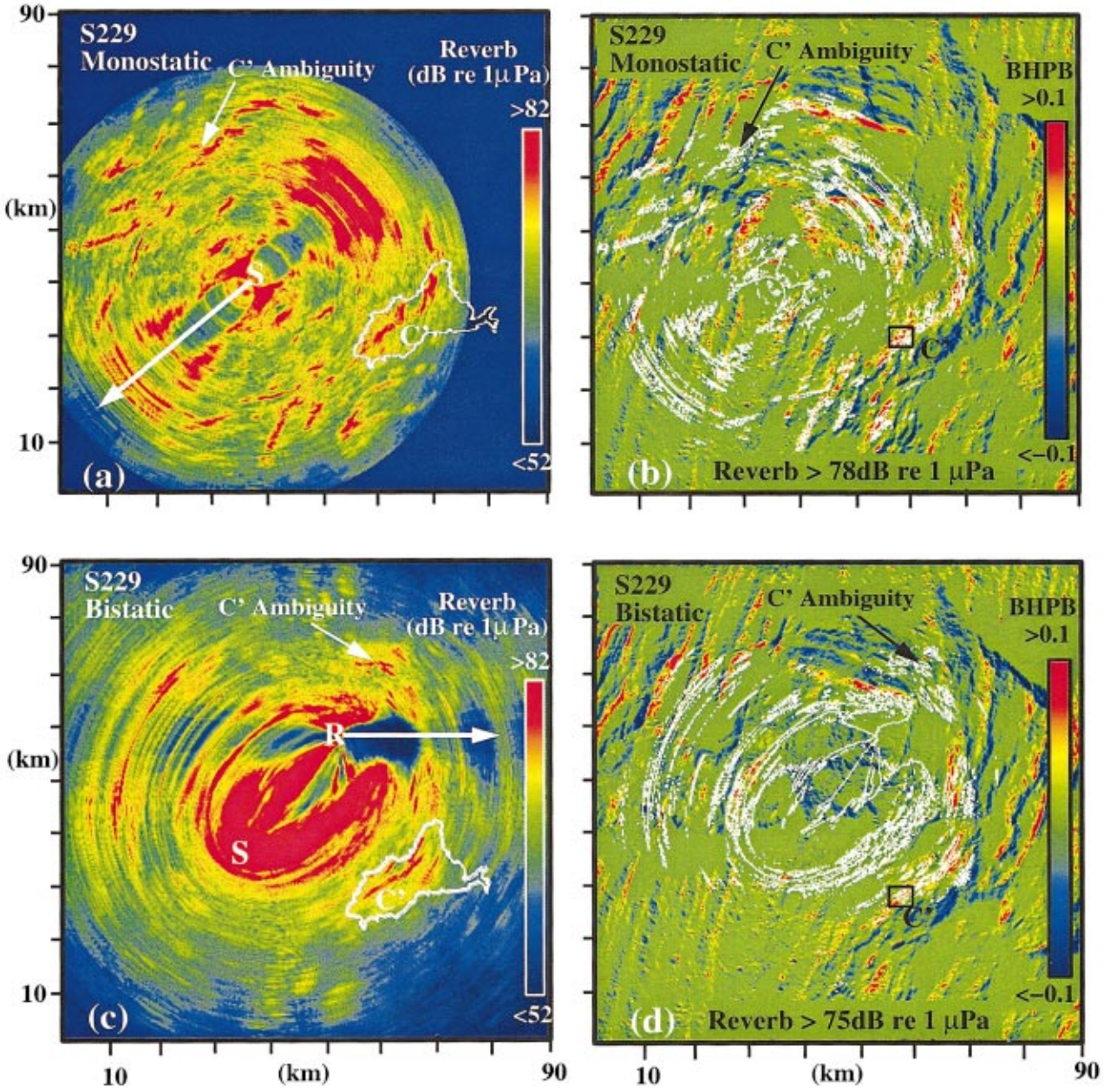


FIG. 14. Wide-area images of monostatic and bistatic reverberation measured for 200–255-Hz LFM S229. (a) Monostatic reverberation chart showing symmetry about the array axis for CORY heading at 228°. (b) Contours of high-level backscatter, overlain on the BHBP, coregister with major scarps on C′ facing the source–receiver. (c) Bistatic reverberation chart showing asymmetry about the ALLIANCE's array heading at 91°. (d) Contours of high-level backscatter overlain on the BHPB. The SW box is shown in black in (b) and (d).

order of the 5.6-dB standard deviation of speckle noise.[6] Since prominent echoes returned from the SW corner scarp are tens of dB higher than returns from neighboring scarps, the large-scale structure of the scarp can be imaged deterministically in both monostatic and bistatic receptions. There is also a possibility that the scarp area at C′ might contain some small-scale features, which are under-resolved at 200-m resolution, responsible for the scattered echo patterns observed in the bistatic charts. We have found no distinct correlation, however, between fine-scale scarp structure (<200-m scale) and fine-scale structure in prominent returns from the scarps in the B′ high-resolution study.[3] Nor should one be expected, since the expected variations of modeled reverberation along the B′ scarps are on the order of the 5.6-dB speckle noise standard deviation. Signal-dependent noise is therefore believed to be the most probable cause for the lack of fine-scale correlation between measured and modeled reverberation.

D. Biazimuthal scattering distribution of the C′ scarp

Curves of the mean measured and modeled reverberation levels, computed at 200-m resolution over the SW box of C′, $\langle R(x,y|\Omega_i,\Omega_r)\rangle_{A_{C'}}$ and $\langle R_M(x,y|\Omega_i,\Omega_r)\rangle_{A_{C'}}$, are plotted as a function of receiver azimuth Ω_r in Figs. 21(a)

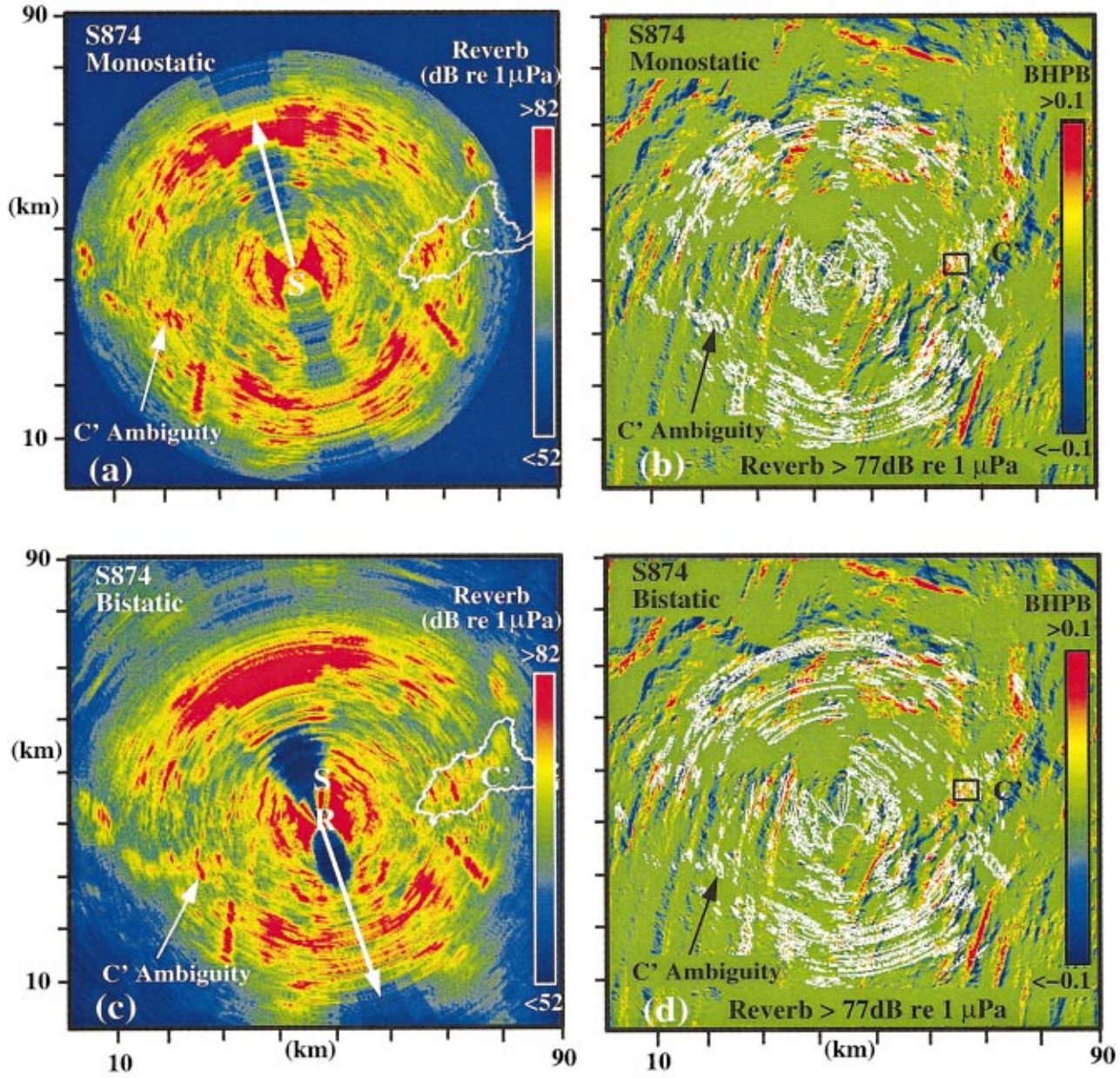


FIG. 15. Wide-area images of monostatic and bistatic reverberation measured for 200–255-Hz LFM S874. (a) Monostatic reverberation chart showing symmetry about the array axis for CORY heading at 345°. (b) Contours of high-level backscatter, overlain on the BHBP, coregister with major scarps on C′ facing the source–receiver. (c) Bistatic reverberation chart showing circular symmetry about the ALLIANCE's array heading at 162°. (d) Contours of high-level backscatter overlain on the BHPB.

and (b), along with their standard deviations $\sigma_{A_{C'}}\{R(x,y)\}$ and $\sigma_{A_{C'}}\{R_M(x,y)\}$. The subscript $A_{C'}$, which follows the same notation used in the B′ study, indicates that the measured and modeled reverberation are averaged over an area A, namely the SW corner scarp of C′. A full biazimuthal description of these parameters, with respect to source and receiver azimuths, can be regained by referring to the source–receiver location pairs as shown in Fig. 12.

The mean measured reverberation curve, in Fig. 21(a), shows a remarkably constant behavior across the receiver azimuths with standard deviation of roughly 5 dB. Compared to the B′ reverberation curves, the average level of the C′ curve is found to be lower by roughly 2 dB. In Fig. 21(b), the mean modeled reverberation curve displays relatively constant behavior, except at the extreme receiver azimuths, $\Omega_r < -60°$, where a roll-off of more than 10 dB is observed. This behavior can be explained by examining the surface projection and two-way TL terms.

Figure 22(a) illustrates the mean surface projection $\langle C_i(x,y|\Omega_i,\Omega_r)+C_r(x,y|\Omega_i,\Omega_r)\rangle_{A_{C'}}$ plotted as a function of receiver azimuth Ω_r. A convex dependence is observed with standard deviations of roughly 6 dB across the receiver azimuths. The mean value peaks at −15 dB near the origin, and gradually rolls off to roughly −20 dB towards the two extremes. This convex behavior is expected since the SW box is comprised of a scarp that faces the center of Easternstar. Thus, the receiver azimuth at 0°, which corresponds to the RV CORY's location at the center of Easternstar, yields a

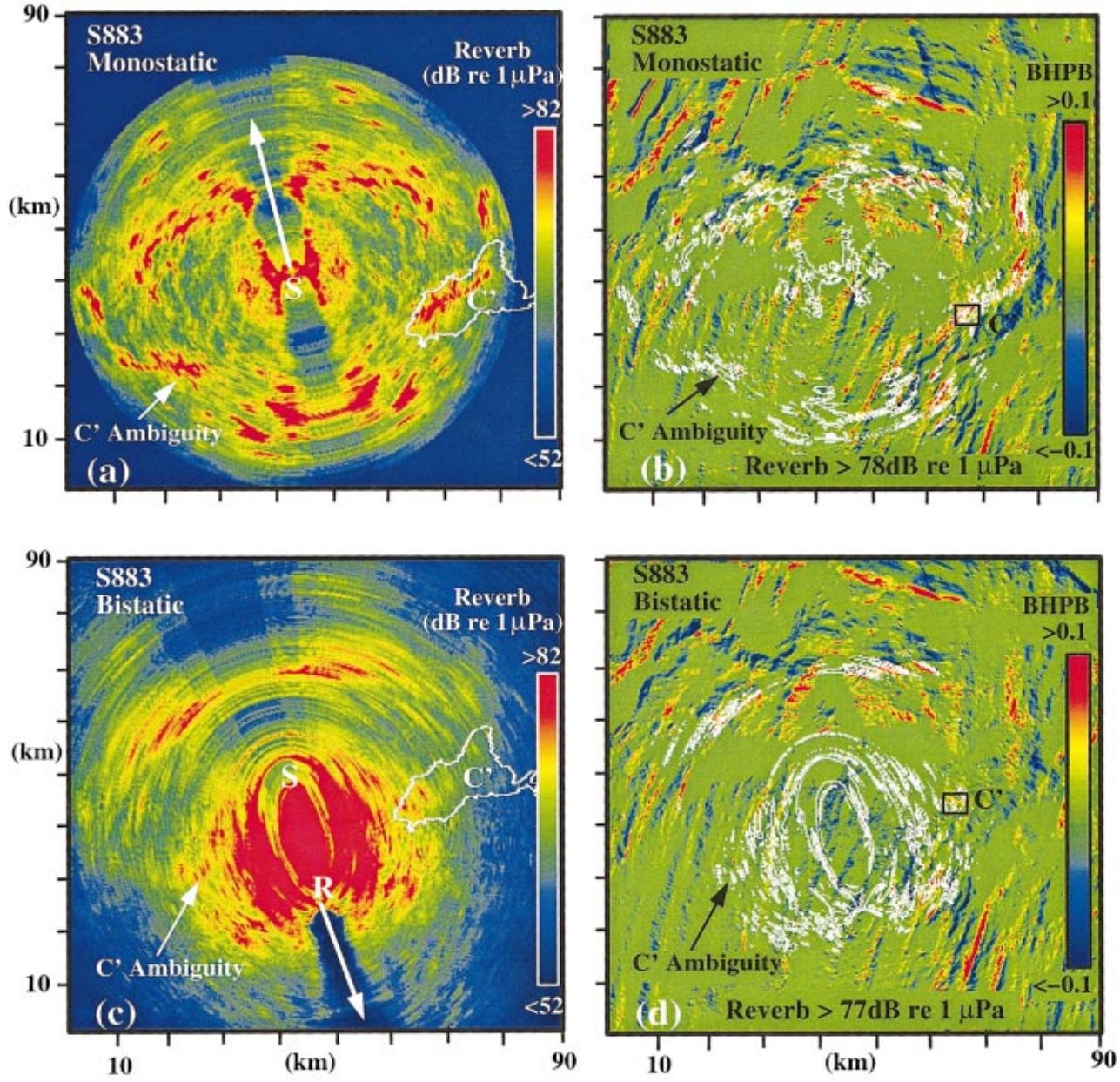


FIG. 16. Wide-area images of monostatic and bistatic reverberation measured for 200–255-Hz LFM S883. (a) Monostatic reverberation chart showing symmetry about the array axis for CORY heading at 346°. (b) Contours of high-level backscatter, overlain on the BHBP, coregister with major scarps on C′ facing the source–receiver. (c) Bistatic reverberation chart showing elliptical symmetry about the ALLIANCE's array heading at 163°. (d) Contours of high-level backscatter overlaid on the BHPB.

higher surface projection than the extreme receiver azimuths.

Figure 22(b) shows the two-way transmission loss $\langle \mathrm{TL}_i(x,y|\Omega_i,\Omega_r)+\mathrm{TL}_r(x,y|\Omega_i,\Omega_r)\rangle_{A_{C'}}$ plotted as a function of receiver azimuth Ω_r. The mean two-way TL curve is relatively flat across the receiver azimuths, except for $\Omega_r < -60°$ where the curve rises up by more than 10 dB. Large two-way TL occurs at this azimuthal extreme because the upper elevation of the SW box apparently lies in the shadow zone of the source's main beam according to the 200-m sampled bathymetry. Lower elevations of the SW box, however, are well insonified by the source's main beam and have low TL. Consequently, a wide spread of TL occurs across these extreme azimuths which leads to large standard deviations in the two-way TL curve for $\Omega_r < -60°$. For $\Omega_r > -60°$, the entire SW box enjoys main-beam insonification and low TL. The modeled reverberation over the C′ scarp has also displayed similar characteristics as the measured reverberation, except for $\Omega_r < -60°$ where our model predicts a stronger shadowing effect than found in the data. This effect is likely due to the use of what is probably highly undersampled bathymetry in our modeling at C′.

The mean strength of the biazimuthal scattering distribution estimated over the C′ scarp, $\langle \hat{\tilde{F}}(x,y|\Omega_i,\Omega_r)\rangle_{A_{C'}}$, is plotted as a function of receiver azimuth Ω_r along with its standard deviation $\sigma_{A_{C'}}\{\hat{\tilde{F}}(x,y)\}$ in Fig. 23. The curve dis-

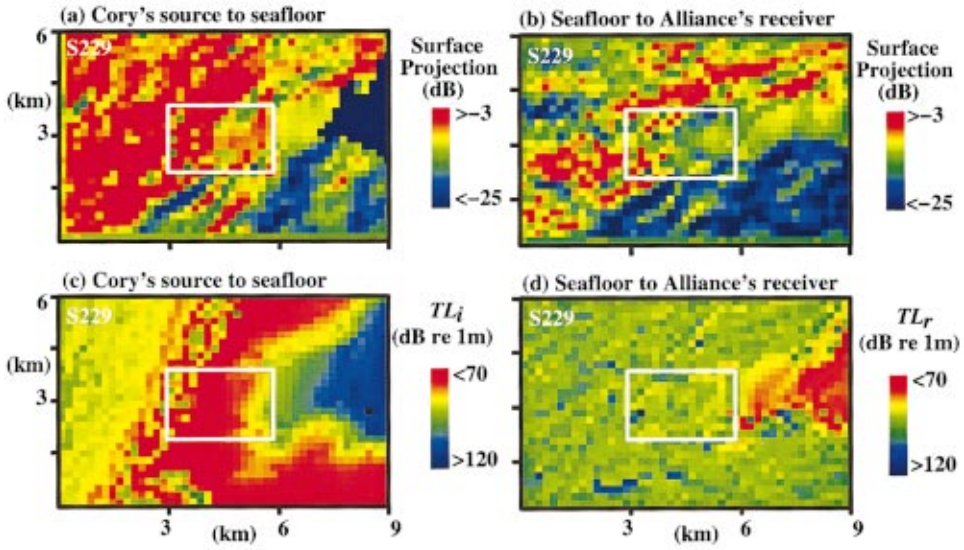


FIG. 17. Surface projection and one-way transmission loss charts computed over the SW corner of C′ for S229 transmission, including the SW box. (a) Surface projection $C_i = 10\log(\cos\theta_i)$ of bathymetry from source to seafloor using ray trace. (b) Surface projection $C_r = 10\log(\cos\theta_r)$ of bathymetry from seafloor to receiver using ray trace. (c) Transmission loss TL_i from CORY source array to seafloor scattering patch. (d) Transmission loss TL_r from seafloor to ALLIANCE receiver.

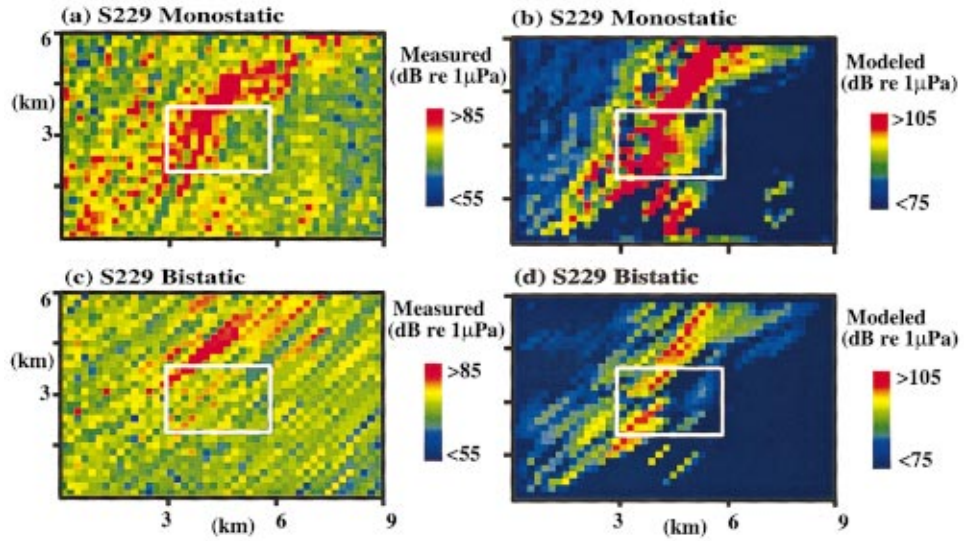


FIG. 18. Charts of the measured and modeled reverberation for S229 over the SW corner of C′ at 200-m resolution including the SW box. (a) Measured monostatic reverberation. (b) Modeled monostatic reverberation. (c) Measured bistatic reverberation. (d) Modeled bistatic reverberation.

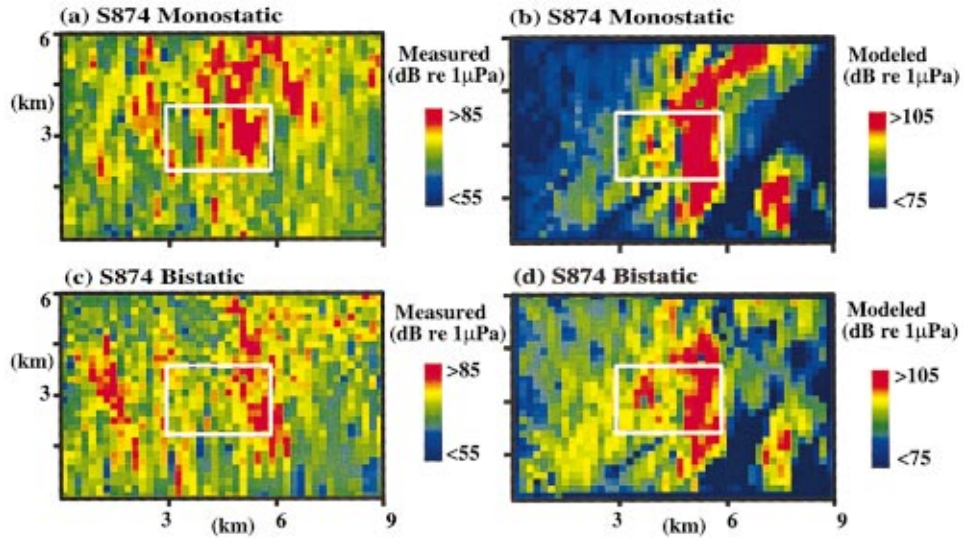


FIG. 19. Charts of the measured and modeled reverberation for S874 over the SW corner of C′ at 200-m resolution including the SW box. (a) Measured monostatic reverberation. (b) Model monostatic reverberation. (c) Measured bistatic reverberation. (d) Modeled bistatic reverberation.

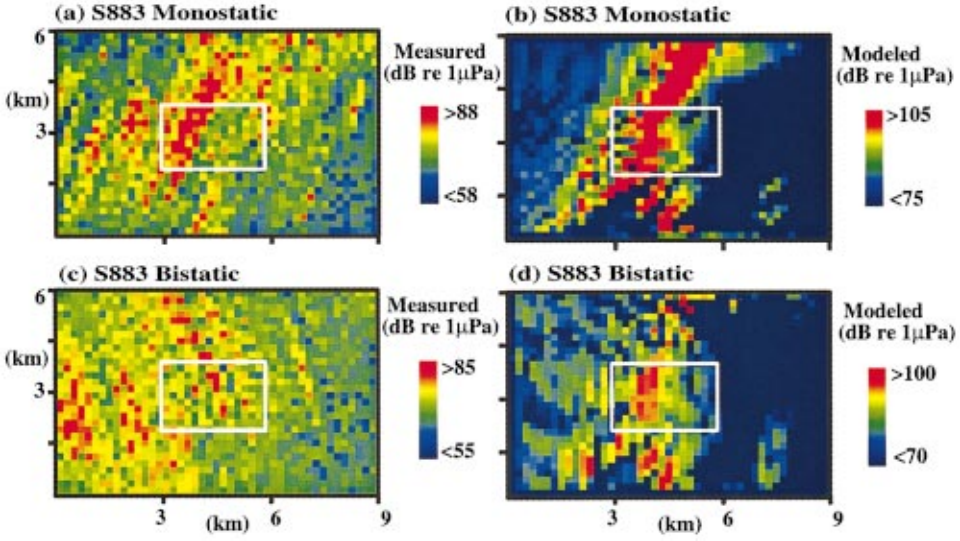


FIG. 20. Charts of the measured and modeled reverberation for S883 over the SW corner of C′ at 200-m resolution including the SW box. (a) Measured monostatic reverberation. (b) Modeled monostatic reverberation. (c) Measured bistatic reverberation. (d) Modeled bistatic reverberation.

plays a relatively constant behavior, except for $\Omega_r < -60°$, where the roll-off observed at these extreme azimuths is primarily due to the shadowing described in the mean TL curves. An average value of roughly -13 dB is obtained by averaging the mean strength of the scattering distribution across the entire $\pm 90°$ range of receiver azimuths. This constant value falls within all error bars.

V. COMPARISON OF BIAZIMUTHAL SCATTERING DISTRIBUTIONS OVER B′ and C′

The biazimuthal scattering distributions as well as measured and modeled reverberation for the two geologically distinct B′ and C′ abyssal hills are compared in this section at 200-m resolution. First, the curves of the mean reverberation level measured across the C′ scarp and the two B′

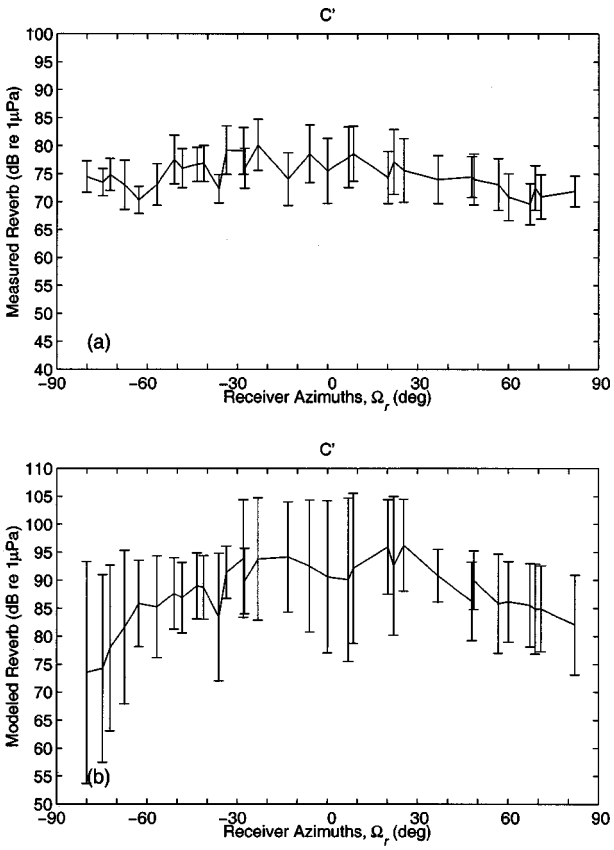


FIG. 21. (a) Mean measured reverberation level $\langle R(x,y|\Omega_i,\Omega_r)\rangle_{A_{C'}}$, and (b) mean modeled reverberation level $\langle R_M(x,y|\Omega_i,\Omega_r)\rangle_{A_{C'}}$ computed over the SW box of C′ as a function of receiver azimuth Ω_r, along with their respective standard deviations, $\sigma_{A_{C'}}\{R(x,y)\}$ and $\sigma_{A_{C'}}\{R_M(x,y)\}$.

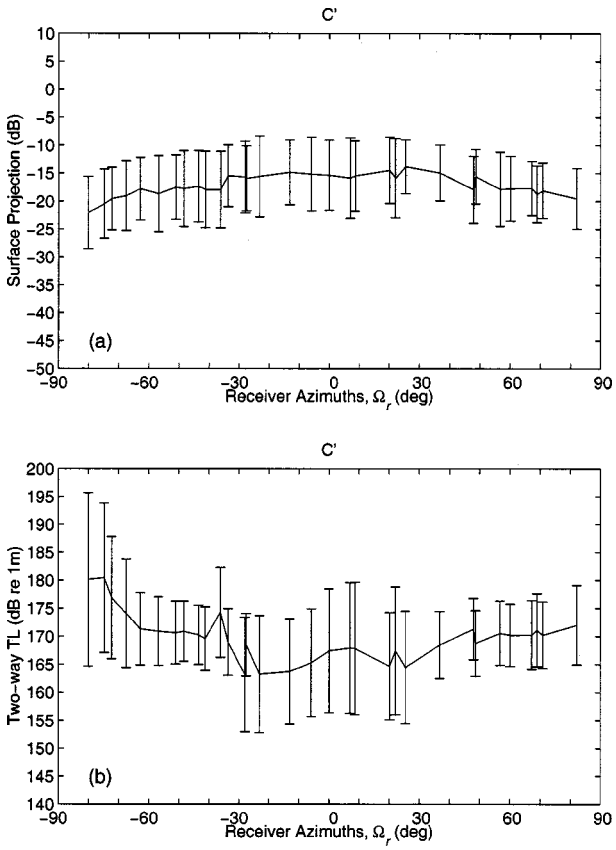


FIG. 22. (a) The mean surface projection $\langle C_i(x,y|\Omega_i,\Omega_r) + C_r(x,y|\Omega_i,\Omega_r)\rangle_{A_{C'}}$, and (b) the mean two-way transmission loss $\langle \mathrm{TL}_i(x,y|\Omega_i,\Omega_r) + \mathrm{TL}_r(x,y|\Omega_i,\Omega_r)\rangle_{A_{C'}}$, computed over the SW box of C′ as a function of receiver azimuth Ω_r.

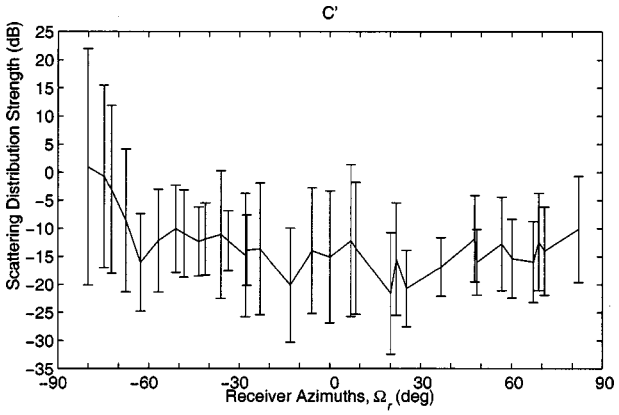


FIG. 23. The mean strength of the biazimuthal scattering distribution estimated over the SW box of C′$\langle \hat{\tilde{F}}(x,y|\Omega_i,\Omega_r)\rangle_{A_{C'}}$ as a function of receiver azimuth Ω_r along with its standard deviation $\sigma_{A_{C'}}\{\hat{\tilde{F}}(x,y)\}$.

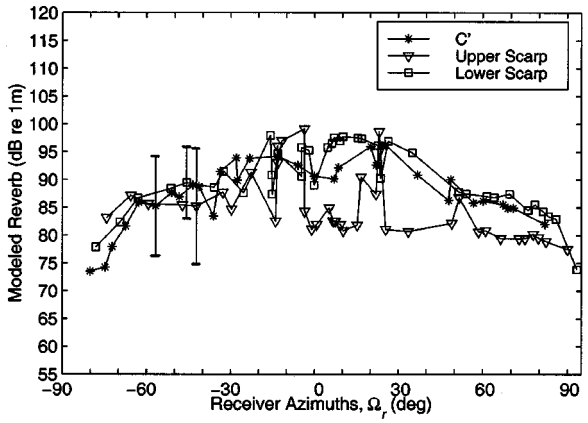


FIG. 25. Comparison of the mean reverberation levels, at 200-m resolution, modeled over the C′ scarp and the two B′ scarps as a function of receiver azimuth Ω_r. Their typical standard deviations are plotted on each curve to illustrate the spread of modeled reverberation.

scarps are plotted as a function of receiver azimuth Ω_r in Fig. 24. A typical standard deviation is plotted on each curve to illustrate the spread of measured reverberation over these scarps. Generally, the mean values of all three curves are relatively uniform across the $\pm 90°$ receiver azimuths. Although the reverberation levels measured over the C′ scarp are occasionally 2–5 dB lower than those of the B′ scarps at some receiver azimuths, these differences lie within the roughly 6-dB standard deviation of all curves. As a result, one may conclude that there is no significant difference between the mean reverberation levels measured over the B′ and C′ scarps at $\frac{1}{2}$ CZ, and that at $\frac{1}{2}$ CZ reverberation measured over the major scarps of these two distinct abyssal hills is homogeneous across nonforward receiver azimuths.

Curves of the mean reverberation level modeled over the C′ scarp and the two B′ scarps, at 200-m resolution, are plotted as a function of receiver azimuth Ω_r in Fig. 25. While the B′ lower scarp and C′ SW scrap show regular convex behavior with peak values within $|\Omega_r|<30°$, the B′ upper scarp exhibits some fluctuations within $|\Omega_r|<30°$, which often exceed 10 dB. As discussed in Ref. 3, these fluctuations occur when the upper scarp falls into the shadow zone of the source main beam's refractive path. A similar shadowing effect is apparently revealed in the C′ curve for $\Omega_r<-60°$, where a 10-dB roll-off is observed. The modeled reverberation curves over the B′ lower scarp and C′ show a better match, for $\Omega_r>-60°$, since these two sites are well insonified by the source's main beam.

The strengths of the biazimuthal scattering distributions, estimated over the C′ scarp and the two B′ scarps, are plotted as a function of receiver azimuth Ω_r in Fig. 26. Large fluctuations for the B′ upper scarp are observed within $|\Omega_r|<30°$ due to shadowing. A constant line at roughly -11 dB can be drawn across the entire set of receiver azimuths for the C′ and B′ curves that still falls within all the error bars of roughly 10 dB and is centrally located when the means of all three curves are taken into account.

VI. DISCUSSION AND CONCLUSIONS

The bistatic scattering characteristics of two geologically distinct abyssal hills located on the western flank of the Mid-Atlantic Ridge, composed of *outside corner* and *inside corner* crust and referred to as B′ and C′, respectively, are experimentally compared. The levels of bistatic reverberation, measured from scarps on the two abyssal hills in bi-

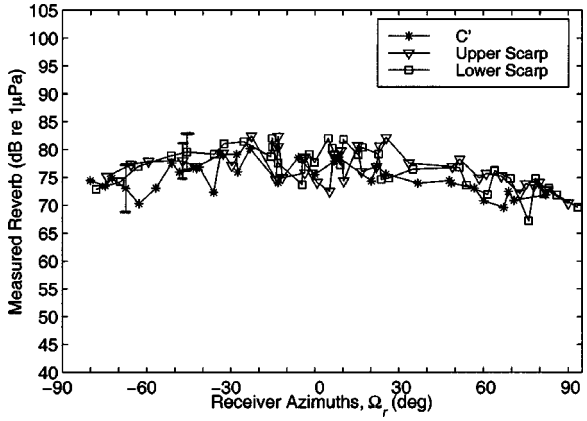


FIG. 24. Comparison of the mean reverberation levels, at 200-m resolution, measured over the C′ scarp and the two B′ scarps as a function of receiver azimuth Ω_r. Their typical standard deviations are plotted on each curve to illustrate the spread of measured reverberation.

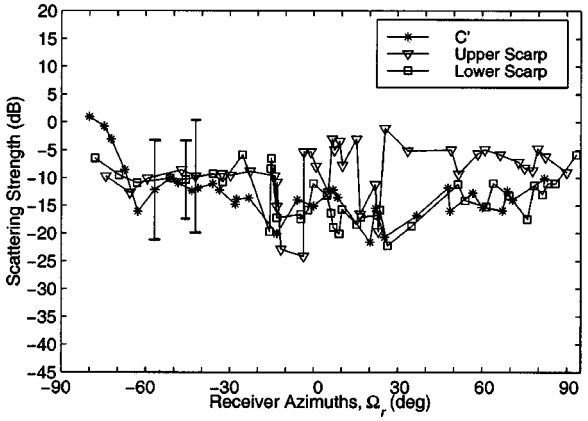


FIG. 26. Comparison of the mean biazimuthal scattering distribution strengths, at 200-m resolution, estimated over the C′ scarp and the two B′ scarps as a function of receiver azimuth Ω_r. Their typical standard deviations are plotted on each curve to illustrate the spread of mean strengths.

static experiments from $\frac{1}{2}$ CZ stand-off, exhibit nearly identical, constant azimuthal dependencies. The mean biazimuthal scattering distributions of scarps on the two abyssal hills are also found to exhibit nearly identical and constant azimuthal dependencies with mean strength equal to -11 dB *when estimated from supporting bathymetry sampled at* 200-m *intervals*. Higher-resolution supporting bathymetry, only available at B′ and not at C′, sampled at 5-m intervals reveals that the projected area of the B′ scarps, as seen by refracted rays traveling from source to bistatic receiver at $\frac{1}{2}$ CZ, is significantly undersampled with the 200-m sampled bathymetry. This undersampling leads to a uniform bias of roughly -6 dB in the level of modeled bistatic reverbation from the B′ scarps and, consequently, a uniform bias of $+6$ dB in the strength of the mean biazimuthal scattering distributions of the B′ scarps. The strength of the mean biazimuthal scattering distributions of the B′ scarps is more accurately given by the constant -17 dB ± 8 dB when estimated from the high-resolution bathymetry sampled at 5-m intervals. A general conclusion is that the use of bathymetry that undersamples the projected area of the seafloor within the resolution footprint of the towed-array system can lead to significant overestimates in the strength of seafloor scattering.

It is significant that although the introduction of 5-m sampled bathymetry revealed the B′ scarps to be comprised of a highly nonplanar network of canyons and gullies ($\sim$200-m scale) that is not resolved in the 200-m sampled bathymetry, the azimuthal dependence of the mean scattering distributions of the scarps remained constant when estimated with the two different bathymetric samplings and only differed by a constant 6-dB offset. The explanation is related to the fact that the sonar system resolution footprint, with cross range extent of roughly 1 km at $\frac{1}{2}$ CZ, typically averages over many canyons and gullies and that the mean biazimuthal scattering distribution averages over a large number of footprints on a given scarp. Delicate, coherent directional differences in scattering due to small-scale structures, such as canyons and gullies, then average out while the overall expected level is still sensitive to the mean projected area within a given resolution footprint since, even on average, this is a first-order quantity directly proportional to the total flux received from the patch.

We expect the same argument to apply to scattering from the C′ scarp, and so hypothesize that the mean biazimuthal scattering distribution will still have constant strength, but will be reduced by a fixed positive offset if the 200-m sampled bathymetry at C′ underestimates the projected area of seafloor sites within the system resolution footprint. Our expectation is that the 200-m sampled bathymetry does indeed underestimate these projected areas.

We then conclude that long-range reverberation from prominent geological features of the Mid-Atlantic Ridge, and likely other midocean ridges, can be adequately modeled as having Lambertian scattering characteristics. We hypothesize that the albedo of $\pi/10^{1.7}$, measured for the two major scarps on the B′ abyssal hill with a more than adequate bathymetric sampling density, provides a reasonable estimate of the albedo of all abyssal hills comprised of *outside corner* crust and may also provide a good estimate of the albedo of abyssal hills comprised of *inside corner* crust. We take the albedo of $\pi/10^{1.1}$, measured for an abyssal hill comprised of *inside corner* crust from potentially undersampled bathymetry, as an upper bound on albedos of abyssal hills comprised of *inside corner* crust.

APPENDIX

This Appendix provides the equations used for modeling reverberation and estimating the mean biazimuthal scattering distribution strength. Derivations of these equations are given in Sec. 2C of Ref. 3. The resolution footprint, at $\frac{1}{2}$ CZ of 33-km range, occupies an annular sector with dimensions of roughly 50 by 1000 m that are large compared to the mean acoustic wavelength of 6.7 m. The total received field from the sonar footprint can therefore be treated as stochastic and diffuse. The expected intensity radiated from the resolution patch can be charted to the center of that patch by performing the convolution

$$\langle I(x,y)|r_i,r_r)\rangle \approx \int\int_{A_f(x,y|r_i,r_r)} f(\theta_i,\phi_i;\theta_r,\phi_r) w l_i l_r \cos\theta_i \cos\theta_r \, dX\, dY, \tag{A1}$$

where θ and ϕ denote the elevation and azimuth angles with respect to the seafloor's surface at (X,Y), and the subscripts i and r indicates the incident and scattered angles, respectively. The function f is known as the bidirectional scattering distribution function (BSDF) of the surface, which is similar to the concept of bidirectional reflectance distribution function in the modern radiometry.[3,15] Note that the traditional scattering strength in underwater acoustics corresponds to $S=f\cos\theta_i,\cos\theta_r$. When f is an angle-independent constant, it is related to the surface albedo α by $f=\alpha/\pi$, and becomes equivalent to the coefficient μ commonly used in the seafloor scattering-strength estimation. The other factors in Eq. (A1) include the transmission power w, and the transmission factors l_i and l_r. This result can be expressed in decibels as

$$R_{\langle I\rangle}(x,y|r_i,r_r) \approx W + 10\log\left(\int\int_{A_f(x,y|r_i,r_r)} 10^{[F(\theta_i,\phi_i;\theta_r,\phi_r)-\mathrm{TL}_i-\mathrm{TL}_r+C_i+C_r]/10} dX\, dY \right), \tag{A2}$$

with the following set of notations: $R_{\langle I\rangle}$: Reverberation level in dB re 1 μPa, F: Biazimuthal scattering distribution's strength in dB re 1 μPa@1 m, $F=10\log f$, W: Source level in dB re 1 μPa@1 m, $W=10\log w$, TL_i: Transmission loss from source to scattering patch in dB re 1 m, $\mathrm{TL}_i=10\log l_i$, TL_r: Transmission loss from the scattering patch to the receiver in dB re 1 m, $\mathrm{TL}_r=10\log l_r$, C_i: Surface projection in the direction of incident refracted wave in dB, $C_i=10\log(\cos\theta_i)$, C_r: Surface projection in the direction of scattered refracted wave in dB, $C_r=10\log(\cos\theta_r)$, A_f: Sonar resolution footprint area in square meters.

Modeled reverberation is from elemental seafloor patches with $F=0$ that scatter equally in all directions so that

$$R_M(x,y|r_i,r_r) = W+10\log\left(\int\int_{A_f(x,y|r_i,r_r)} 10^{[-\mathrm{TL}_i-\mathrm{TL}_r+C_i+C_r]/10}\,dX\,dY\right). \tag{A3}$$

This leads to a simple linear equation for the mean strength of the scattering distribution over the resolution footprint

$$\overline{F}(x,y|r_i,r_r)=R_{\langle I\rangle}(x,y|r_i r_r)-R_M(x,y|r_i,r_r), \tag{A4}$$

where the dependence on the incident and scattered angles is integrated over the resolution footprint. Finally, the maximum likelihood estimate for $\overline{F}$ is given by

$$\hat{\overline{F}}(x,y|r_i,r_r)=R(x,y|r_i,r_r)-R_M(x,y|r_i,r_r), \tag{A5}$$

which is the difference between reverberation measured with MAE data and modeled with $F=0$.

[1] Acoustic Reverberation Special Research Program, Initial Report, edited by J. Orcutt, Scripps Institution of Oceanography (1993).
[2] N. C. Makris, ''Proposed experiment,'' in Acoustic Reverberation Special Research Program Research Symposium, Woods Hole Oceanographic Institution (1992).
[3] N. C. Makris, C. S. Chia, and L. T. Fialkowski, ''The bi-azimuthal scattering distribution of an abyssal hill,'' J. Acoust. Soc. Am. **106**, 2491–2512 (1999).
[4] N. C. Makris, L. Avelino, and R. Menis, ''Deterministic reverberation from ocean ridges,'' J. Acoust. Soc. Am. **97**, 3547–3574 (1995).
[5] B. E. Tucholke and J. Lin, ''A geological model for the structure of ridge segments in slow spreading ocean crust,'' J. Geophys. Res. **99**, 11937–11958 (1994).
[6] N. C. Makris, ''The effect of saturated transmission scintillation on ocean acoustic intensity measurements,'' J. Acoust. Soc. Am. **100**, 769–783 (1996).
[7] B. E. Tucholke, J. Lin, M. C. Kleinrock, M. A. Tivey, T. B. Reed, J. Goff, and G. E. Jaroslow, ''Segmentation and crustal structure of the western Mid-Atlantic Ridge flank, 25°25′–27° 10′N and 0–29 m.y.,'' J. Geophys. Res. **102**, 10203–10223 (1997).
[8] K. B. Smith, W. S. Hodgkiss, and F. D. Tappert, ''Propagation and analytic issues in the prediction of long-range reverberation,'' J. Acoust. Soc. Am. **99**, 1387–1404 (1996).
[9] A. J. Harding, M. A. H. Hedlin, and J. A. Orcutt, ''Migration of backscatter data from the Mid-Atlantic Ridge,'' J. Acoust. Soc. Am. **103**, 1787–1803 (1998).
[10] C. S. Chia, ''A Comparison of Bistatic Scattering from Two Geologically Distinct Mid-Ocean Ridges,'' S. M. thesis, MIT, Cambridge, MA, 1999.
[11] N. C. Makris and B. Gardner, ''Planned tracks/waypoints, runs 3–9,'' in Acoustics Reverberation Special Research Program Main Acoustics Experiment, Initial Report, Scripps Institution of Oceanography (1993), pp. 65–80.
[12] J. R. Preston, E. Michelozzi, L. Troiano, and R. Hollett, *Cruise Report on RV ALLIANCE cruise MARE 5 July–1 August 1993 SACLANTCEN's Joint Experiment with ONR's ARSRP Group*, Report M-112 (SACLANT Undersea Research Centre, LaSpezia, Italy, 1993).
[13] N. C. Makris and J. M. Berkson, ''Long-range backscatter from the Mid-Atlantic Ridge,'' J. Acoust. Soc. Am. **95**, 1865–1881 (1994).
[14] N. C. Makris, ''Imaging ocean-basin reverberation via inversion,'' J. Acoust. Soc. Am. **94**, 983–993 (1993).
[15] B. K. P. Horn and R. W. Sjoberg, ''Calculating the reflectance map,'' Appl. Opt. **18** (No. 11), 1770–1779 (1979).

Penetration of acoustic waves into rippled sandy seafloors

E. Pouliquen,[a] A. P. Lyons, and N. G. Pace
SACLANTCEN, Viale S. Bartolomeo 400, I-19138 La Spezia, Italy

(Received 22 October 1998; revised 25 January 2000; accepted 24 May 2000)

The Helmholtz–Kirchhoff integral and the Kirchhoff approximation are applied to model the penetration of sound waves into rough sandy seafloors at grazing angles above and below the critical angle. As the seafloor of interest is anisotropic, emphasis is placed on simulating the response from a two-dimensional interface. The analytical development of the method is first presented, followed by numerical examples. Simulations and data acquired at sea are in very good agreement in the 2–15 kHz band [Maguer *et al.*, J. Acoust. Soc. Am. **107**, 1215–1225 (2000)]. The model predicts, in agreement with the 2–15 kHz acoustic data, the contributions due to roughness effects that dominate the evanescent wave component over most of this frequency band. Secondary effects such as coherent (Bragg) influence patterns and the loss of signal coherence with grazing angle or depth are correctly predicted. The model simulations strongly suggest that roughness of the sediment interface is most likely the cause of anomalous sound penetration into the seabed.
[S0001-4966(00)00910-3]

PACS numbers: 43.30.Gv, 43.30.Hw, 43.30.Ma [DLB]

I. INTRODUCTION

Interest in the penetration of acoustic energy into the seafloor below the critical grazing angle has increased with the relatively recent experimental results that sound intensity is often above that predicted by elastic wave theory for smooth seabeds (Maguer *et al.*,[1] Chotiros,[2,3] and Lopes[4]). In these experiments the inclusion of various probable penetration mechanisms such as the evanescent wave or bounded beam effects (Jensen and Schmidt,[5] Williams *et al.*[6]) underestimated the actual level of acoustic energy transmitted across the water–sediment interface. Other mechanisms that may explain these anomalous levels include the Biot slow wave (Chotiros[2]) and roughness scattering (Thorsos *et al.*[7]). Using a broadband parametric system (2–15 kHz), Maguer *et al.*[1] have been able to map, for their experimental conditions, the point in frequency where the dominant mechanisms affecting the pressure field in the sediment changes from the evanescent wave and bounded beam effects to another mechanism, which they attribute to roughness scattering. Although the question of which mechanism, interface roughness, or Biot slow wave, plays a greater role in the transmission of acoustic energy into the seafloor at subcritical grazing angles has not yet been definitively answered, it is becoming increasingly accepted that roughness is, in many cases, an important contributor (Simpson *et al.*,[8] Maguer *et al.*[1]). In this context, the objective of this paper is to find a viable explanation for anomalous penetration through a complex anisotropic rough surface around the critical angle as well as obtaining a relative order of magnitude for the two suspected main mechanisms: the evanescent wave and roughness of the interface.

Roughness scattering models of subcritical penetration of sound into marine sediments have generally used perturbation theory with an isotropic power law spectrum to describe seafloor roughness (Thorsos *et al.*,[7] Jackson *et al.*[9]). Both the assumption of an isotropic power law spectrum and the validity of perturbation theory in a two-dimensional (2-D) plane come into question for many shallow-water, sandy seafloors. In shallow water, the interaction of currents with the water–sediment interface can be very strong, commonly producing relatively large, oriented, quasiperiodic features, i.e., ripply fields. Ripple fields are usually formed in granular deposits with little or no cohesion between particles and are observed in both unidirectional and oscillatory flows. At low sediment transport rates, the wavelength of the ripples is directly proportional to the orbital amplitude of the fluid outside the boundary layer with a constant of proportionality that does not depend greatly on grain size or density. These types of ripples are often called vortex ripples because of the vortex that forms in the lee of each crest and have waveheight-to-wavelength ratios of 0.1–0.25 (Sleath[10]). Once the ripple field is formed it can persist when the fluid velocity falls below the limit for sediment movement. The ripple field structure will remain until it is changed by another mechanism such as bioturbation or another strong wind-wave event. For sandy seafloors in shallow water, these roughness formation mechanisms are such that the isotropic power law is often not an accurate description of the roughness spectrum as it ignores the larger-scale anisotropic and near-monochromatic spectral structure (two to ten cycles/m) that exists. Information on the exact form of the small-scale roughness spectra for shallow water seafloors is usually based on stereo photogrammetric methods (Akal,[11] Briggs,[12] Wheatcroft[13]) with results often presented in the form of a one-dimensional spectrum. One-dimensional analysis is very limiting for the study of anisotropic seafloors, such as those with oriented sand ripples. A power law has been fitted to rippled seafloor roughness spectra taken along and across the ripple orientations by Briggs[14] and the obtained power law parameters were subsequently used in predictions of scattering strength (Jackson *et al.*[15]). A dominant frequency component does not often appear in the data. The reasons for this

[a]Electronic mail: pouliq@saclantc.nato.int

are unclear but one possibility is that the processing steps used in Briggs[14] have the tendency to smear or reduce frequency resolution, either by averaging adjacent lines or by the windowing/prewhitening operations (even though the leakage or bias is reduced at higher frequencies). Although the general high-frequency trend is fitted by a power law, any relatively small bandwidth components (such as ripples) that may be present are obscured in this type of analysis. Additionally, even if smoothing was not a problem, visual identification of distinct frequency components is hindered by the log–log plots that are normally used to display roughness spectra. As will be shown in this article, it is the lower-frequency, narrow-bandwidth components that are particularly important in calculations of high-frequency subcritical penetration on rippled sand seafloors.

In addition to considering the anisotropic character of water–sediment interfaces, the three-dimensional nature of the seafloor must also be considered. There is an inconsistency between scattering results from 2-D scattering models (e.g., using a 1-D profile) and 3-D scattering models (using 2-D surfaces) due principally to the fact that 2-D models do not consider ''out of plane'' scattering. Results showing this effect have been presented by Yang *et al.*,[16] which were based on the small slope approximation for 2-D isotropic rough Dirichlet surfaces. The small slope approximation can be considered among the most accurate theories to date, Yang's calculations bring suspicion on any 2-D scattering model (even if they agree very well with a 2-D ''exact solution''). If there is a mismatch between 2-D and 3-D models, it is fair to strongly question the accuracy of 2-D models in their attempt to match reality. This last statement is also valid for the case of penetration through a rough interface separating two fluids. To account for the problem mentioned above, the model presented in this paper has a three-dimensional approach based on the Kirchhoff approximation (KA). This approximation has strong limitations but makes the implementation of a 3-D time domain penetration model possible and relatively easy. Also, at high acoustic frequencies and for relatively smooth seafloors with correlation lengths that are long enough or values of root mean square (RMS) roughness that are small enough, the KA can be used with confidence at grazing angles as low as 15° to essentially provide an ''order of magnitude'' calculation of the effect of roughness on penetration. Other advantages of the KA approach are that the model is intuitive and accommodating to the evanescent wave and finite beam effects. The three-dimensional distribution of features on the seafloor are fully included in the model by summing contributions from a seafloor surface that is generated using realistic roughness statistics. As it is the quasiperiodic, lower-frequency structure that will be most important for the method presented here, a noncentered Gaussian spectrum is used for constructing realistic seafloor surfaces. The spectrum is also filtered to exclude the nonrelevant lowest stochastic scales. Correct beam patterns can be included in order to accurately predict the effect of the plane wave spectrum on subcritical penetration.

In Sec. II, a discussion of realistic rippled seafloors is presented with their spectra and some examples of generated surfaces. In Sec. II we also outline the development of the penetration model. In Sec. III we present simulation results obtained using this model for various geometric and physical parameters. These simulations illustrate the ability of the model to predict the influence of roughness, the evanescent wave, and the bounded beam effect, as well as the grazing angle and depth dependence of coherent (Bragg) interference effects to the penetration. In the last section we summarize the conclusions of the article.

II. SEAFLOOR INTERFACE PENETRATION MODELING

The penetration model presented here is based on the Kirchhoff approximation, also called the ''tangent plane approximation.'' This approximation assumes that the reflection coefficient and transmission coefficient that are derived for an infinite plane wave encountering an infinite plane surface can be used at every point of a rough surface. Briefly, its validity depends on the character of the rough surface in question and whether or not this surface has sharp features. Because sandy seafloors are often gently smooth, we will hypothesize that the Kirchhoff approximation is reasonable at a low grazing angle, even for frequencies as low as 2–3 kHz to provide a qualitative and a rough quantitative indication of penetration levels into the seabed. The basic principle of the method consists of summing coherently the elementary pressure signals from all the surface points.

A. Modeling the seafloor interface

A filtered Gaussian spectrum was chosen to describe the rippled seafloor interface. This choice is motivated by the fact that the high-frequency region does not have a critical importance, as emphasis is on large-scale roughness in the present model. This would not be the case, however, if the approach was based on perturbation theory, which would require accurate estimation of the high-frequency part of the spectrum as well. The chosen power spectrum $W_s(\mathbf{K})$ is able to account for either isotropic or nonisotropic features and the general form is expressed analytically as

$$W_s(\mathbf{K}) = Q_s(\mathbf{K}, \mathbf{K_c}) + Q_s(\mathbf{K}, -\mathbf{K_c}). \tag{1}$$

In Eq. (1), the spatial frequency $\mathbf{K_c} = (K_{xc}, K_{yc})$ defines the average ripple wavelength and Q_s, which has the properties of a joint probability distribution function, is given by

$$\begin{aligned} Q_s(\mathbf{K}, \mathbf{K_c}) = {} & \frac{l_x l_y \eta^2}{4\pi\sqrt{1-\rho^2}} e - [l_x^2 (K_x - K_{xc})^2 + l_y^2 (K_y - K_{yc})^2 \\ & - 2\rho l_x l_y (K_x - K_{xc})(K_y - K_{yc})]/ \\ & 2(1-\rho^2), \quad \text{if } |\mathbf{K}| > K_{\text{hp}}, \\ = {} & 0, \quad \text{if } |\mathbf{K}| \leqslant K_{\text{hp}}, \end{aligned} \tag{2}$$

where K_x, K_y are the components of wave number, l_x, l_y are the components of the correlation length. The spatial frequency K_{hp} sets the limit of a high-pass filter that removes very large-scale features and that is necessary for typical seafloors for which low-frequency content is absent (e.g., on a very large scale, the seafloor is more or less flat). η is a factor that allows the RMS roughness variance σ_s^2 to be

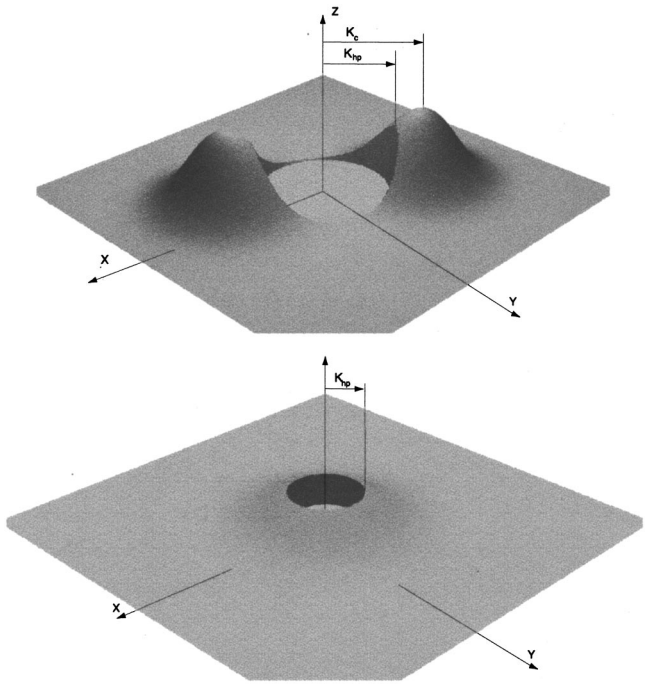


FIG. 1. Nonisotropic filtered Gaussian spectrum (top) and isotropic filtered Gaussian spectrum (bottom).

specified independently of K_{hp} or $\mathbf{K_c}$. The parameter ρ is the correlation coefficient between the X and the Y directions and for the sake of simplicity will be set to 0 in this study. Depending on the amplitude of K_c, l_x, and l_y, the seafloor can have stronger or weaker directional features. The power spectrum W_s is normalized so that

$$\int_s W_s(\mathbf{K})\,d\mathbf{K}=\sigma_s^2. \tag{3}$$

If $\mathbf{K_c}=(\mathbf{0},\mathbf{0})$ and $l_x=l_y=l$, the seafloor has no directional features and is isotropic and Eq. (2) reduces to the expression

$$W_s(K_x,K_y)=\frac{l^2\eta^2}{2\pi}e^{-l^2K^2/2}, \quad \text{if } |\mathbf{K}|>K_{\mathrm{hp}},$$
$$=0, \quad \text{if } |\mathbf{K}|\leqslant K_{\mathrm{hp}}. \tag{4}$$

Figure 1 shows examples of both the anisotropic and isotropic spectra produced with Eqs. (1), (2), and (5). For the isotropic case, η can be related to the RMS roughness by the expression

$$\eta^2=\sigma_s^2\cdot e^{K_{hp}^2 l^2/2}. \tag{5}$$

The spectrum [Eqs. (1) and (2)] is utilized for generating realistic stochastic realizations of the water–sediment interfaces based on the Fourier synthesis method.[17] For each seafloor realization, a stochastic surface height power spectrum is generated and converted to an amplitude field in the spatial domain using a Fourier transform. Each realization has the expected first and second moments. The generated power spectrum consists of a predefined roughness power spectrum [for example, W_s of Eq. (1)] in which part of the phase and the amplitude is randomized. Each random power spectrum produces a unique height field configuration that can be utilized for computing a single realization of the pressure field at a given location below the generated interface by the method outlined in the next section. Figure 2 is an example of a seafloor interface height field realization produced by the above method. This surface displays a ripple field that is skewed by 30° with respect to the observer.

B. Helmholtz–Kirchhoff formulation

The source is located far from the water–sediment interface so that one can assume an effective plane wave insonifying a rough surface with a grazing angle θ_g. In this study, multiple scattering from the water–sediment interface, scattering from the sediment volume and shear waves are not considered. The seafloor volume is considered homogeneous and propagation loss in the sediment is only due to the compressional wave attenuation (β). Within the restriction on the validity of the Kirchhoff approximation, for a fluid–fluid case, the boundary conditions of pressure and normal velocity continuity across the interface lead to the following relations:

$$p_t(\mathbf{R})=[1+\mathcal{R}_{12}(\mathbf{R_1},\mathbf{1},\mathbf{n}]p_i(\mathbf{R}), \tag{6}$$

and

$$\frac{\partial}{\partial n}p_t(\mathbf{R})=\frac{\overline{\rho_2}}{\overline{\rho_1}}[1-\mathcal{R}_{12}(\mathbf{R_1},\mathbf{n})]\frac{\partial}{\partial n}p_i(\mathbf{R}). \tag{7}$$

The pressure field $p_i(\mathbf{R})$ is the incoming pressure field in the water at the local point $\mathbf{R}$ (Fig. 4) and $p_t(\mathbf{R})$ represents the pressure field transmitted through the water–sediment interface. The terms $\mathcal{R}_{12}(\mathbf{R_1},\mathbf{n})$ and $(1+\mathcal{R}_{12}(\mathbf{R_1},\mathbf{n}))$ are the local water–sediment, plane wave reflection, and transmission coefficients at the point $\mathbf{R}$, respectively. The terms $\overline{\rho_2}$ and $\overline{\rho_1}$ are the averaged water and sediment densities. $p_i(\mathbf{R})$ can be expressed in the farfield for a slowly varying directivity pattern as

$$p_i(\mathbf{R})=p_0R_0D_i(\mathbf{R_1},f)G_i(\mathbf{R_1}), \tag{8}$$

where p_0 is the transmitted pressure level, the term R_0 is a unit of distance, and $D_i(\mathbf{R_1},f)$ is the farfield directivity gain in the direction $\mathbf{R_1}=(\mathbf{P}-\mathbf{R})$. This directivity term is assumed to depend on frequency and to be slowly varying with the angle. The term $G_i(\mathbf{R_1})$ is the Green's function for the incident pressure field,

$$G_i(\mathbf{R_1})=\frac{e^{-j\overline{\mathbf{k}_1}\cdot\mathbf{R_1}}}{|\mathbf{R_1}|}=\frac{e^{-j\overline{k}_1R_1}}{R_1}, \tag{9}$$

where $R_1=|\mathbf{R_1}|$. Clay and Medwin[18] propose a formulation of the Helmholtz–Kirchhoff integral for the case of sound penetration through a rough surface when a receiver is located at $\mathbf{R}'$ (see Fig. 3). Here, we consider a differential equation of the H–K formulation by looking at the elementary contribution from an element of surface $d\mathbf{S_R}$ located from the point $\mathbf{R}$ on the interface to the hydrophone located in $\mathbf{R}'$ below this surface. We assume that the elementary surface $d\mathbf{S_R}$ located by $\mathbf{R}$ is contained in the contributing surface points S_t:

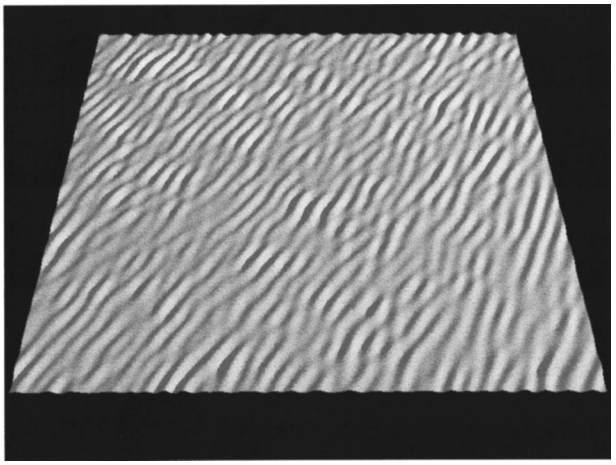


FIG. 2. The 3-D display of a portion of the seafloor realization $z(\mathbf{R})$ used in the following simulations. The size of the displayed surface is 6 m×6 m, the RMS roughness σ_s^2 is set to 2.5 cm, $K_c=31.5$ rad/m, and $K_{\text{hp}}=10$ rad/m.

$$dp(\mathbf{R}')=\frac{1}{4\pi}\left[p_t(\mathbf{R})\frac{\partial}{\partial\mathbf{n}}G_t(\mathbf{R_2})-G_t(\mathbf{R_2})\frac{\partial}{\partial\mathbf{n}}p_t(\mathbf{R})\right]d\mathbf{S_R}. \tag{10}$$

The operator $\partial/\partial\mathbf{n}$ is a spatial derivative in a direction normal to the surface and $\mathbf{R_2}=|\mathbf{R}-\mathbf{R}'|$. The Green's function $G(\mathbf{R_2})$ of Eq. (10) represents the propagation from the point $\mathbf{R}$ to the point $\mathbf{R}'$. It can be expressed as

$$G_t(\mathbf{R},\mathbf{R}')=\frac{e^{-j\overline{\mathbf{k_2}}\cdot\mathbf{R_2}}}{|\mathbf{R_2}|}=\frac{e^{-j\overline{k_1}\overline{n_2}R_2}}{R_2}. \tag{11}$$

Combining Eqs. (7) and (8), one obtains

$$\frac{\partial}{\partial\mathbf{n}}p_t(\mathbf{R})=p_0R_0\frac{\overline{\rho_2}}{\overline{\rho_1}}[1-\mathcal{R}_{12}(\mathbf{R_1},\mathbf{n})]D_i(\mathbf{R_1},f)\frac{\partial}{\partial\mathbf{n}}G_i(\mathbf{R_1}). \tag{12}$$

A closer look at $(\partial/\partial\mathbf{n})G_i(\mathbf{R_1})$ shows that

$$\frac{\partial}{\partial\mathbf{n}}G_i(\mathbf{R_1})=\boldsymbol{\nabla}\cdot\mathbf{n}[G_i(\mathbf{R_1})]=\left(\frac{\mathbf{R_1}\cdot\mathbf{n}}{|\mathbf{R_1}|}\right)\frac{\partial}{\partial R_1}\left(\frac{e^{-j\overline{k_1}R_1}}{R_1}\right). \tag{13}$$

It is assumed that $\overline{k_1}R_1\gg1$ (a source far from the interface). We can then approximate (13) by

$$\begin{aligned}\frac{\partial}{\partial\mathbf{n}}G_i(\mathbf{R_1})&=-\left(\frac{\mathbf{R_1}\cdot\mathbf{n}}{|\mathbf{R_1}|}\right)j\overline{k_1}G_i(\mathbf{R_1})\\&=j\ \cos(\gamma(\mathbf{R_1},\mathbf{n}))\overline{k_1}G_i(\mathbf{R_1}),\end{aligned} \tag{14}$$

with $(\mathbf{R_1}\cdot\mathbf{n})/|\mathbf{R_1}|=-\cos(\gamma(\mathbf{R_1},\mathbf{n}))$ and $\gamma(\mathbf{R_1},\mathbf{n})$ being the angle between the incident direction and the normal vector $\mathbf{n}$ to the surface at $\mathbf{R}$. The second term in the bracket of the right-hand side of (10) then equals

$$\begin{aligned}-G_t(\mathbf{R_2})\frac{\partial}{\partial n}p_t(\mathbf{R})=&-j\ \cos(\gamma(\mathbf{R_1},\mathbf{n}))\overline{k_1}\frac{\overline{\rho_2}}{\overline{\rho_1}}\\&\times[1-\mathcal{R}_{12}(\mathbf{R_1},\mathbf{n})]p_i(\mathbf{R})G_t(\mathbf{R_2}).\end{aligned} \tag{15}$$

In the same way,

$$p_t(\mathbf{R})\frac{\partial}{\partial n}G_t(\mathbf{R_2})=-\left(\frac{\mathbf{R_2}\cdot\mathbf{n}}{|\mathbf{R_2}|}\right)\left(jk_2+\frac{1}{R_2}\right)p_t(\mathbf{R})G_t(\mathbf{R_2}). \tag{16}$$

Assuming that $\overline{k_2}R_2\gg1$, which is the case if we consider the hydrophone sufficiently far from the water–sand interface, we obtain

$$\begin{aligned}p_t(\mathbf{R})\frac{\partial}{\partial n}G_t(\mathbf{R_2})=&-j\ \cos(\gamma'(\mathbf{R_2},-\mathbf{n}))\overline{k_1}\overline{n_2}\\&\times[1+\mathcal{R}_{12}(\mathbf{R_1},\mathbf{n})]p_i(\mathbf{R})G_t(\mathbf{R_2}).\end{aligned} \tag{17}$$

Then, combining (15), (17), and (10),

$$\begin{aligned}dp(\mathbf{R}')=&-\frac{j\overline{k_1}}{4\pi}p_0R_0D_i(\mathbf{R_1},f)G_i(\mathbf{R_1})G_t(\mathbf{R_2})\\&\times\left[\frac{\overline{\rho_2}}{\overline{\rho_1}}[1-\mathcal{R}_{12}(\mathbf{R_1},\mathbf{n})]\cos(\gamma(\mathbf{R_1},\mathbf{n}))\right.\\&\left.+\overline{n_2}[1+\mathcal{R}_{12}(\mathbf{R_1},\mathbf{n})]\cos(\gamma'(\mathbf{R_2},-\mathbf{n}))\right]d\mathbf{S_R}.\end{aligned} \tag{18}$$

Here, we assume that the reflection coefficient is constant with frequency over the considered bandwidth. The average index of refraction $\overline{n_2}$ has to be considered as a complex entity as the sediment is assumed to be dissipative:

$$\overline{n_2}=\overline{n_2'}-j\overline{n_2''}. \tag{19}$$

For a non-CW signal of spectrum $E(f)$, grouping together all the frequency-dependent terms, Eq. (18) becomes

$$\begin{aligned}dp(\mathbf{R}',f)=&-\frac{j}{2\overline{c_1}R_1R_2}p_0R_0\\&\times\left[\frac{\overline{\rho_2}}{\overline{\rho_1}}[1-\mathcal{R}_{12}(\mathbf{R_1},\mathbf{n})]\cos(\gamma\mathbf{R_1},\mathbf{n}))\right.\\&\left.+\overline{n_2}[1+\mathcal{R}_{12}(\mathbf{R_1},\mathbf{n})]\cos(\gamma'(\mathbf{R_2},-\mathbf{n}))\right]\\&\times[f\cdot E(f)D_i(\mathbf{R_1},f)e^{-j\overline{k_1}(R_1+\overline{n_2'}R_2)}\\&\times e^{-\overline{k_1}\overline{n_2''}R_2}]d\mathbf{S_R},\end{aligned} \tag{20}$$

and if we consider that the attenuation β in the seafloor is frequency dependent, we have

$$\beta=\alpha|f|=2\overline{k_1}\overline{n_2''}, \tag{21}$$

where α is the attenuation coefficient that is assumed to be fixed at a given sediment porosity. Transforming to the time domain, using the inverse Fourier transform, we obtain

$$dp(\mathbf{R}',t)=\mathcal{F}^{-1}\{dp(s)(\mathbf{P}',\mathbf{f})\} \tag{22}$$

$$=-\frac{j}{2\overline{c_1}R_1R_2}p_0R_0 \times\left[\frac{\overline{\rho_2}}{\overline{\rho_1}}[1-\mathcal{R}_{12}(\mathbf{R_1},\mathbf{n})]\cos(\gamma(\mathbf{R_1},\mathbf{n})) +\overline{n_2}[1+\mathcal{R}_{12}(\mathbf{R_1},\mathbf{n})]\cos(\gamma'(\mathbf{R_2},-\mathbf{n}))\right] \times\mathcal{F}^{-1}[f.E(f)D_i(\mathbf{R_1},f) \times e^{-j\overline{k_1}(R_1+\overline{n_2'}R_2)}e^{-\alpha|f|R_2/2}]d\mathbf{S_R}. \tag{23}$$

Considering the following relation from Ref. 19:

$$\mathcal{F}^{-1}[f\cdot E(f)e^{-j\overline{k_1}(R_1+\overline{n_2'}R_2)}]=\frac{1}{2\pi j}e'\left(t-\frac{R_1+\overline{n_2'}R_2}{\overline{c_1}}\right), \tag{24}$$

Eq. (24) can be expressed as

$$dp(\mathbf{R}',\mathbf{t})=\mathcal{F}^{-1}[dp(s)(\mathbf{P}',\mathbf{f})] \tag{25}$$

$$=\frac{p_0R_0}{4\pi\overline{c_1}R_1R_2} \times\left[\frac{\overline{\rho_2}}{\overline{\rho_1}}[1-\mathcal{R}_{12}(\mathbf{R_1},\mathbf{n})]\cos(\gamma(\mathbf{R_1},\mathbf{n})) +\overline{n_2}[1+\mathcal{R}_{12}(\mathbf{R_1},\mathbf{n})]\cos(\gamma'(\mathbf{R_2},-\mathbf{n}))\right] \times e'\left(t-\frac{R_1+\overline{n_2'}R_2}{\overline{c_1}}\right)*\mathcal{F}^{-1}[e^{-\alpha|f|R_2/2}]* \times\mathcal{F}^{-1}[D_i(\mathbf{R_1},f)]d\mathbf{S_R}. \tag{26}$$

Computing the total pressure field received at the hydrophone consists of summing all the elementary pressures $dp_{(S)}(\mathbf{P}',\mathbf{t})$ over the insonified surface (S_t):

$$p(\mathbf{R}',t)=\int_{(S_t)}dp(\mathbf{R}',\mathbf{t}). \tag{27}$$

The computation of $p(\mathbf{R}',t)$ of Eq. (27) is based on the stochastic realization of seafloor interface height fields $z(\mathbf{R})$ presented in Sec. II A.

III. SIMULATION RESULTS

A. Definition of the penetration ratio

It is necessary to define entities that quantify the penetration of sound waves into the sediment. For a transmitted pulse of spectrum $E(f)$, the intensity received at the hydrophone buried into the sediment at frequency f can be expressed as

$$E_{bh}(f)=\frac{1}{\overline{\rho_2}\overline{c_2}}p(\mathbf{R}',f)p^*(\mathbf{R}',f). \tag{28}$$

The received intensity just above the surface is given by the expression

$$E_{ph}(f)=\frac{1}{\overline{\rho_1}\overline{c_1}}\left(\frac{p_0^2R_0^2}{R_1^2}D_i^2(\mathbf{R}_d)E^*(f)E(f)\right). \tag{29}$$

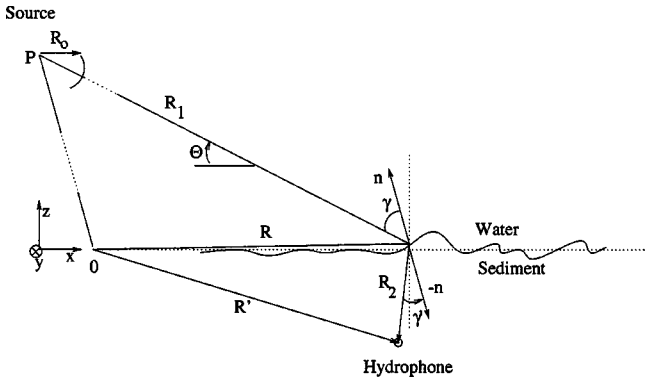


FIG. 3. Measurement geometry for simulations.

The vector $\mathbf{R_d}$ will be picked as being directed toward the surface point for which the directivity pattern is maximum, so that $D_i(\mathbf{R_d})=1$. We define the penetration ratio as

$$\mathrm{PR}(f)=10\log_{10}(E_{bh}(f)/E_{ph}(f))+10\log_{10}\left(\frac{\overline{\rho_2}\overline{c_2}}{\overline{\rho_1}\overline{c_1}}\right). \tag{30}$$

The definition of PR(f) is not an intensity ratio. It corresponds to the transfer function of the water–sand interface. In order to obtain an intensity ratio, the term $10\log_{10}(\overline{\rho_2}\overline{c_2}/\overline{\rho_1}\overline{c_1})$ can be removed from Eq. (30), which results in subtracting about 6 dB from the PR(f) in the case of a sandy seabed. Finally, the total penetration ratio over the whole bandwidth is defined as follows:

$$\mathrm{TPR}=10\log_{10}\left[\int_{-\infty}^{\infty}\mathrm{PR}(f)df\right]. \tag{31}$$

The values of the TPR are strongly sensitive to the hydrophone position. As observed experimentally at sea and with the model simulation, the position of the hydrophone greatly affects E_{bh}, PR(f) and TPR. Thus, reference to a nominal grazing angle θ_g without specifying the exact location of the hydrophone below the seafloor interface may lead to unexpected penetration levels. Sensitivity to the hydrophone position is discussed later in this article.

At sea experiments performed by Maguer *et al.*[1] in the Gulf of Biodola (island of Elba, Italy) used a constant source height (see Fig. 3) with respect to the average seafloor level. This source is positioned on a tower that can be moved horizontally on a rail in only one horizontal direction. As the range between the source and the hydrophone changes, the grazing angle also changes that affects the size of the insonified seafloor patch and also influences the amount of sound penetration. The sensitivity of penetration levels to system geometry emphasizes the need to take geometric parameters into account during sea experiments as well as in simulations. This is particularly crucial around and below the critical angle, where penetration levels vary strongly with angle.

B. Examples of penetration above and below critical angle

The transmitted signal used in the simulation is a wideband pulse centered around 8 kHz. Figure 4 shows the signal and its corresponding power spectrum. This wideband signal,

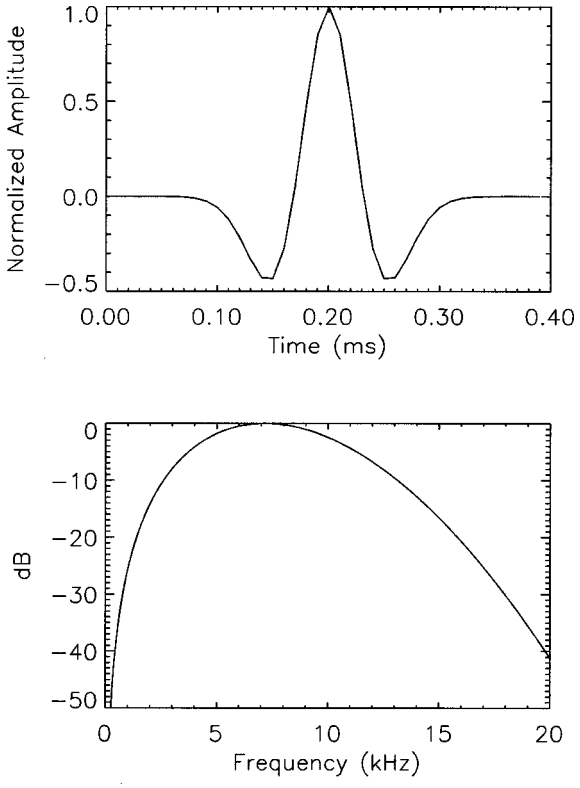


FIG. 4. Ricker signal $e(t)$ used in the simulation (top) and its corresponding power spectrum (bottom).

however, has little energy below 2 kHz and above 15 kHz and, as a consequence, apart from the problem of validity of the present model as frequency decreases, simulated penetration ratios may be inaccurate below 2 kHz. Figure 5 presents model predictions of the amount of intensity received at 8 kHz by the hydrophone located 30 cm underneath the center of the seafloor surface. A directivity pattern close to the one of the parametric source used by Maguer[1] for a 8 kHz Ricker pulse is used here. A theoretical square Gaussian source of width ($2\theta_{3\ \mathrm{dB}}=6°$) is used. $\overline{c_1}=1530$ m/s, $\overline{c_2}=1685$ m/s, $\overline{\rho_2}=1.8$ g/cm^3, $\alpha=0.8$ dB/m/kHz. The energy pattern shows that most of the energy comes from the area directly above the hydrophone. In these simulations, it can be seen that the amount of contributing areas decreases as the grazing angle decreases. The dominant contribution to the total received energy is from the facets oriented toward the source (e.g., facets that present a higher local grazing angle), whereas facets oriented away from the source contribute poorly to penetration.

Figure 6 presents a series of simulated signals received in the sediment above and below the critical grazing angle. As in the at-sea penetration experiment of Maguer,[1] the water sound speed is taken as 1530 m/s and the sediment sound speed as 1685 m/s that corresponds to a critical angle of 24.8°. The generated surface used for these simulations is the same as presented in Fig. 5 with varying RMS roughness. The hydrophone is located 30 cm below the intersection of the center of the beam and the mean level of the surface [Fig. 6 (top)]. An interesting but not unexpected result is that the received signals are mostly coherent above and even at the critical angle (i.e., the received signal resembles the transmitted pulse shape). As the grazing angle decreases, the coherence is degraded but, as low as 20°, the initial part of the received signals remains quite close to the transmitted pulse shape. Below 20°, the coherence deteriorates because the rough character of the surface starts playing a dominant role. It becomes almost impossible to distinguish a shape similar to the transmitted pulse in the received signal at 15°. The bottom plot of Fig. 6 presents the same series of simulated signals for a hydrophone located 60 cm below the water–sand interface. As for a hydrophone located 30 cm below the interface, the coherence is good above the critical angle, while below 20°, the first part of the signal is even more strongly attenuated than at 30 cm.

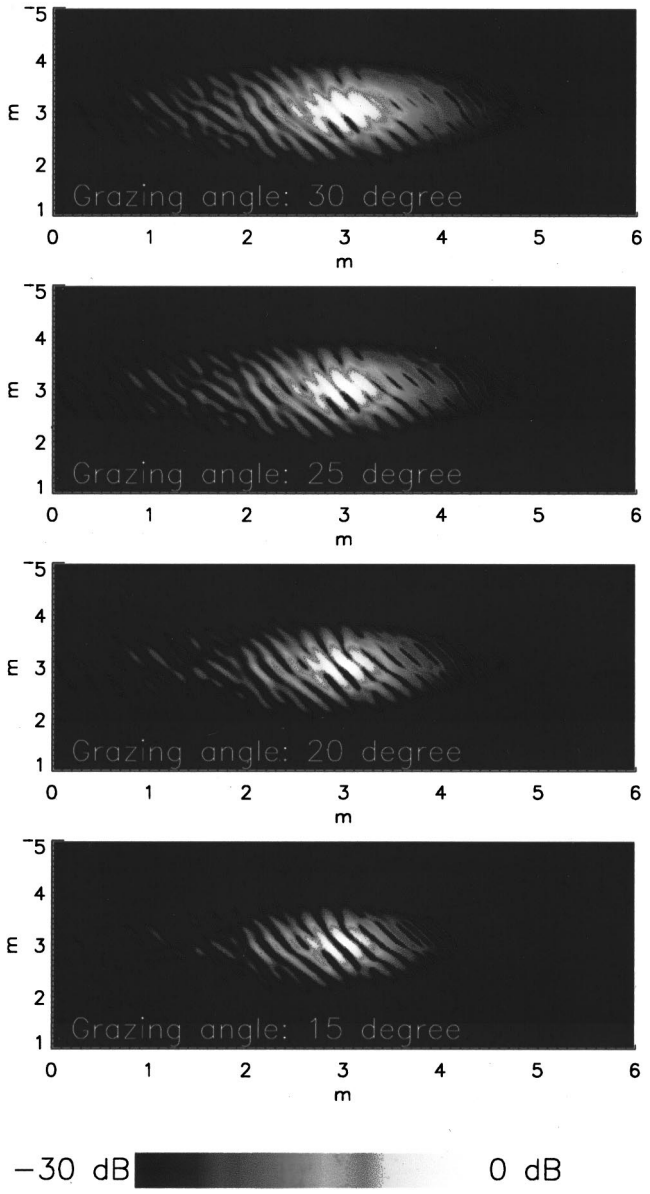


FIG. 5. Energy contributions from the surface of Fig. 2 at four different grazing angles for transmission of the signal of Fig. 4. The source is located on the left and the beam is pointing toward the center of each image. The hydrophone burial depth is 30 cm. From top to bottom: the distance from the source $R_0=16$ m, $\theta_g=30°$; $R_0=18.93$ m, $\theta_g=25°$; $R_0=23.39$ m, $\theta_g=20°$; $R_0=30.91$ m, $\theta_g=15°$. $K_c=31.5$ rad/m, $\sigma_s=1$ cm.

Figure 7 shows penetration ratios corresponding to the signals displayed in Fig. 6. The penetration ratios change

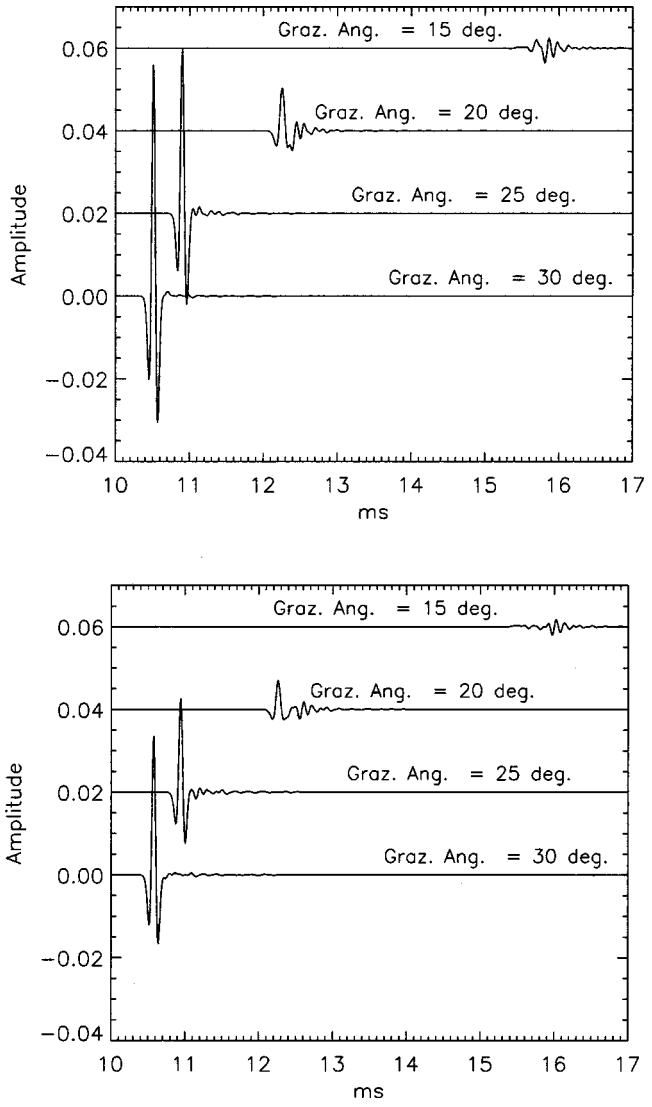


FIG. 6. Example of received pressure signals $p(\mathbf{R}'t)$ versus time at various grazing angles. Input parameters are the same as Fig. 5. The amplitude factor is arbitrary but kept constant for the four signals. The hydrophone is located 30 cm (the top figure) and 60 cm (the bottom figure) into the seafloor.

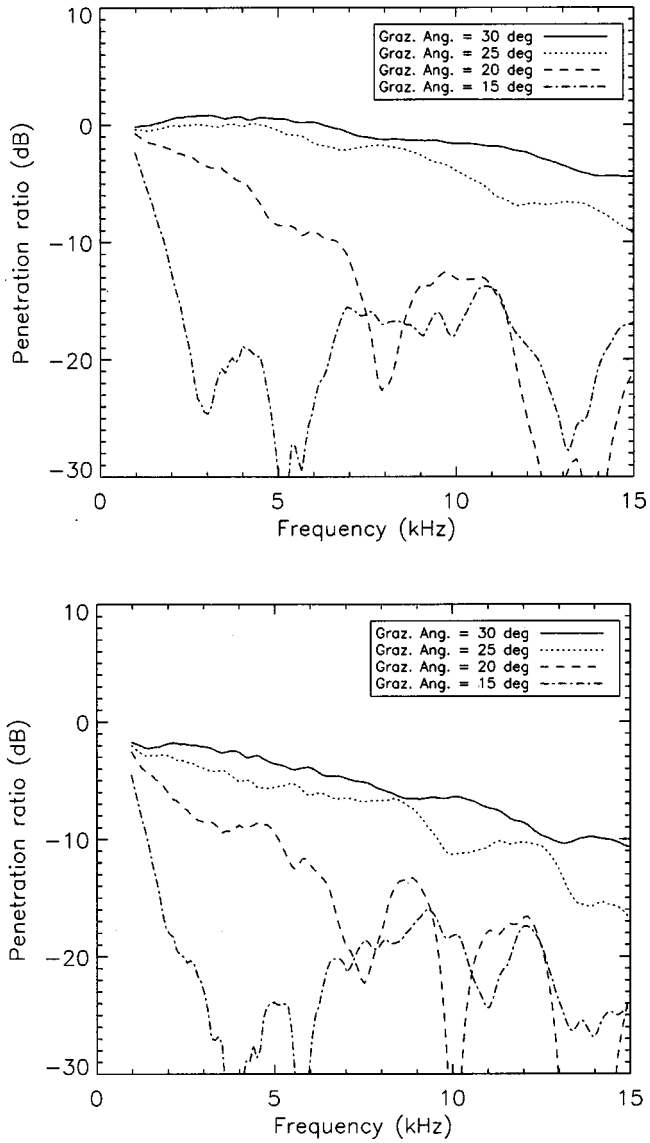


FIG. 7. Example of the penetration ratio PR(f) at various grazing angles. The present penetration ratios correspond to the signals of Fig. 4, for a hydrophone buried 30 cm (top) and 60 cm (bottom) into the seabed.

significantly from 30° to 15°. A quasiconstant level of PR(f) is seen at 30°, which reveals a relatively weak dependence of the penetration on roughness and the expected dependence on attenuation. As the grazing angle decreases, the penetration ratio begins to display an interference structure, which does not appear to be located randomly. As the grazing angle decreases, the minima and maxima are shifted down in frequency, which is in accordance with Bragg theory, as the surface used here contains periodic features. The average correlation length is centered around 20 cm and is skewed by 30° away from the Y axis, which corresponds to an equivalent correlation length of 23 cm in the X direction. Classical Bragg theory computation on the location of minima and maxima is not consistent, however, between the case at 30 cm and the case at 60 cm. Considering the sole interference mechanism produced in the water medium, the locations of the minima and maxima present in Fig. 7 for a burial depth of 60 cm are not in accordance with those for a burial depth of 30 cm. This suggests that the interference phenomenon is not only occurring in the upper medium but also inside the seafloor and that the hydrophone position is a key factor in the construction of interference fringes; for a given location $\mathbf{R}'$ of the hydrophone, there exists a set of frequencies for which constructive and destructive interference will occur. This phenomenon is not easily quantifiable experimentally, but an important difference in the spectral signature of PR(f) has been observed at sea for hydrophones located 50 cm apart.[1] Figures 8 and 9 illustrate this phenomenon. In the top plot of Fig. 8, at a 30° grazing angle (i.e., above critical), the interference patterns change slowly because the roughness effect is dominated by classical refraction mechanisms. As the hydrophone position increases in the X direction, penetration increases because the hydrophone is closer to the main direction of the refracted rays into the sediment (i.e., close to horizontal). In the bottom plot of Fig. 8, below the critical angle at 20°, the penetration ratio is more sensitive to the hydrophone position. This is because below the critical graz-

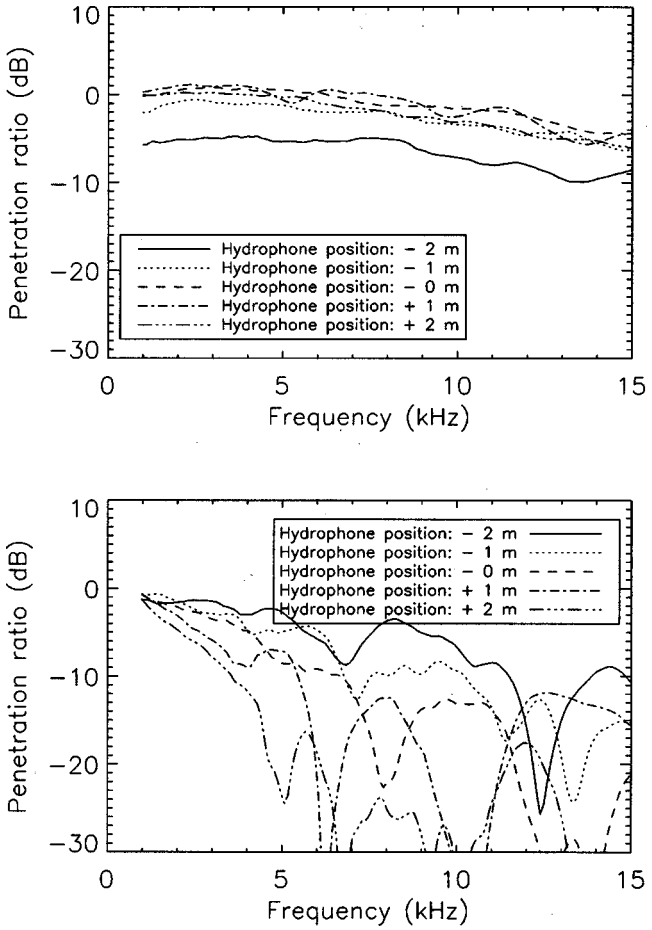


FIG. 8. Example of the variability of the penetration ratio PR(f) for various horizontal positions of the hydrophone. The hydrophone is located in the plane *OXZ*. Hydrophone position $\mathbf{R}' = (x,0,-30$ cm). Top: $\theta_g = 30°$. Bottom: $\theta_g = 20°$. All other geometric and geoacoustic parameters are the same as in the previous figures.

ing angle, the evanescent wave is only significant in the lower-frequency region, leaving the roughness as the only effect above 4–5 kHz in this case, where the importance of the hydrophone position is particularly critical. Figure 9 shows the decay of the TPR below critical as the hydrophone moves away from the source. This decay can be explained by

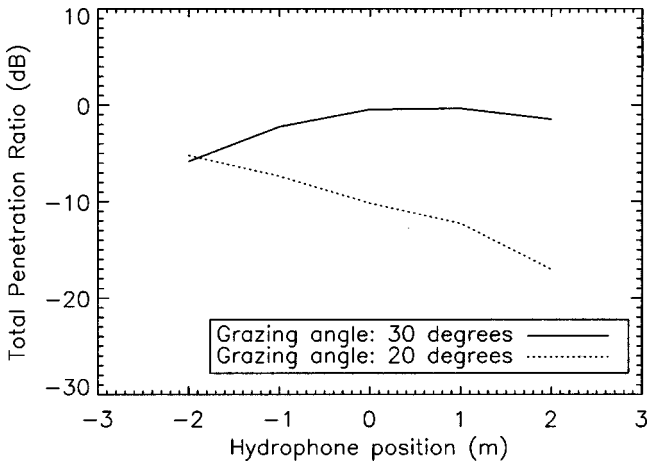


FIG. 9. Total penetration ratio TPR(f) versus the hydrophone horizontal position at two distinct grazing angles (30° and 20°).

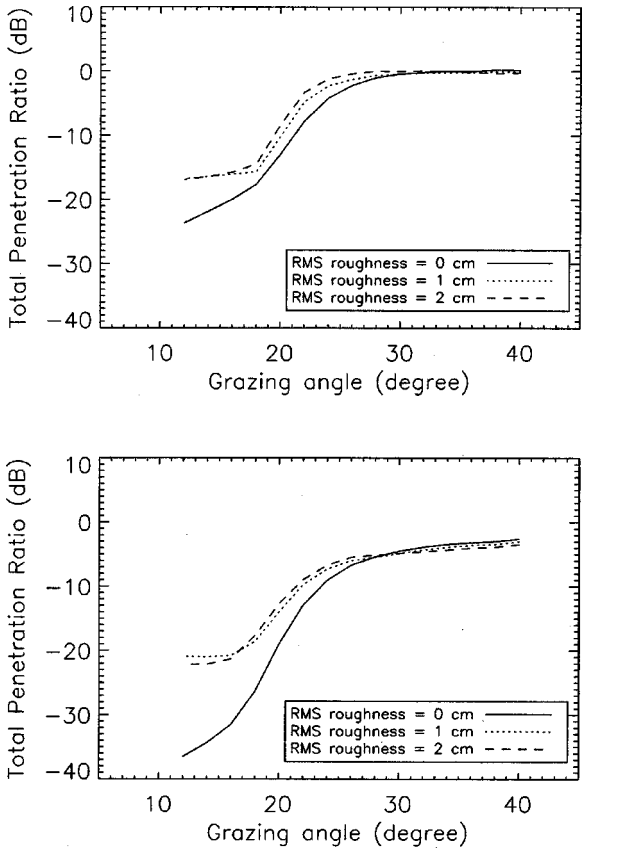


FIG. 10. Total penetration ratio (TPR) versus the grazing angle for various roughnesses and two hydrophone burial depths from cases presented in Figs. 6 and 7.

an increasing distance (e.g., higher attenuation loss) between the hydrophone and the insonified patch. This confirms the importance of the hydrophone location relative to the seafloor height field as well as the particular realization of that seafloor height field. This also suggests that a larger beam aperture would increase the penetration ratio over the whole beamwidth.

C. Parametrization with RMS roughness

The effect of roughness on the total penetration ratio both 30 and 60 cm below the interface is shown in Fig. 10. In

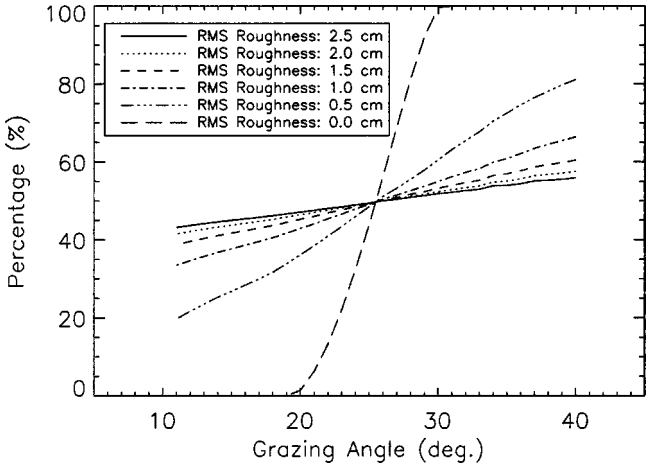


FIG. 11. Percentage of local facets on the seabed that are above critical ($\theta_c = 24.8°$) versus the nominal grazing angle. The geometric and geoacoustic conditions source are the same as the simulation figures.

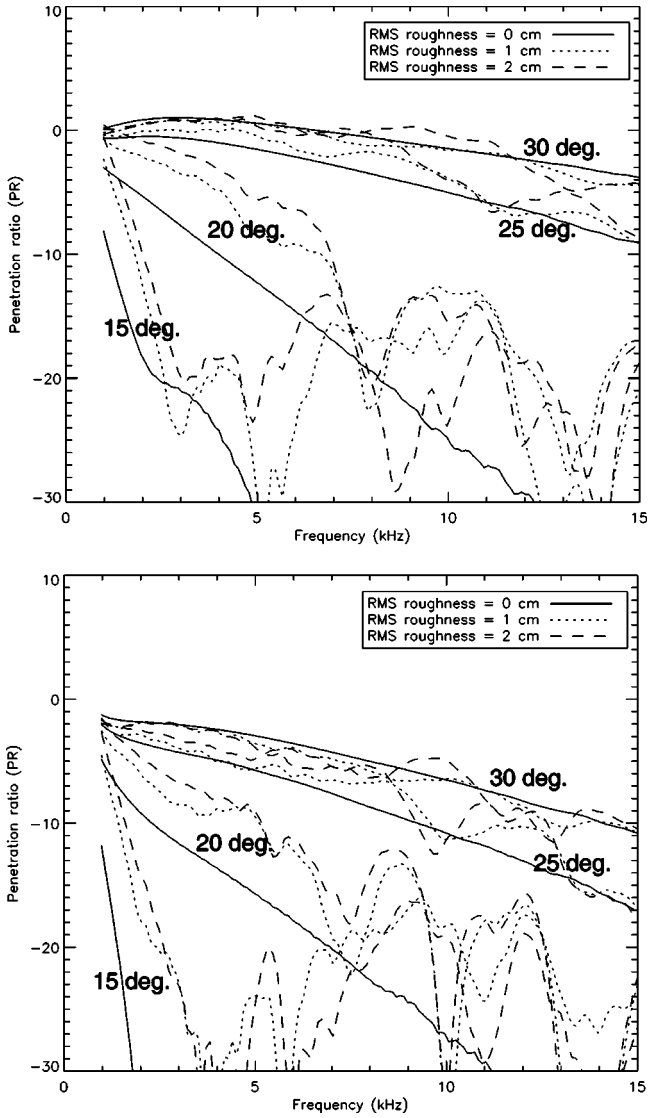


FIG. 12. Penetration ratio PR(f) for various grazing angles and various roughnesses (0, 1, 2 cm RMS). The hydrophone is buried 30 cm (the top figure) and 60 cm (the bottom figure). The geometric and geoacoustic conditions are the same as the previous simulations.

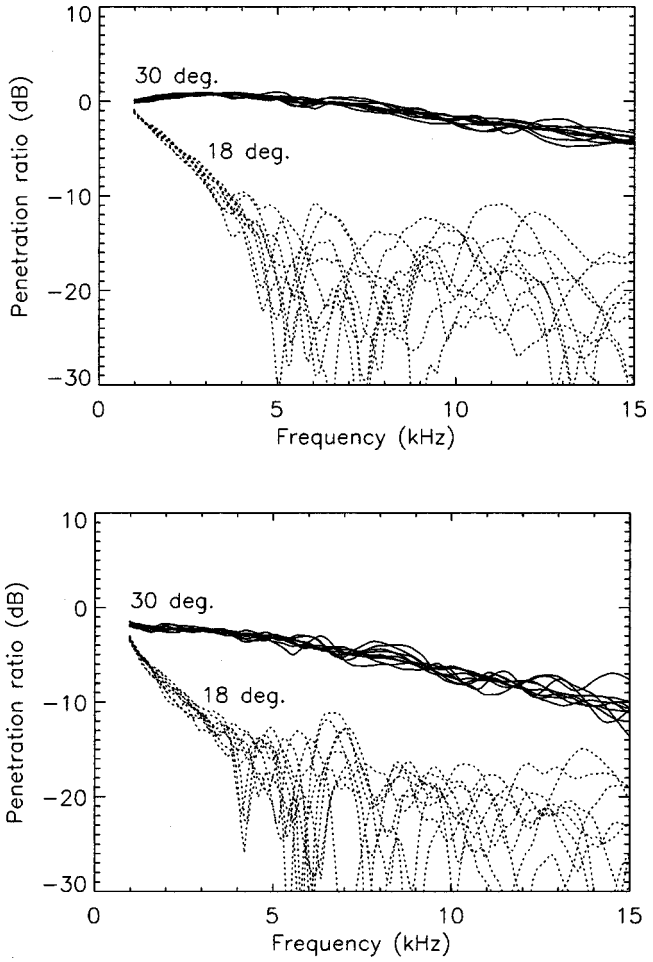


FIG. 13. Example of the sensitivity of the penetration ratio PR(f) for ten realizations of the seafloor having the same statistics: 30 cm burial depth (top), 60 cm burial depth (bottom).

contrast to the case of a flat seafloor, for which, according to classical theory, penetration decreases regularly as the grazing angle decreases, simulations of TPR on a rough surface show that seafloor roughness plays a significant role in the penetration level realized above *and* below the critical angle. Above the nominal critical angle, because of refraction, roughness tends to slightly reduce the penetration level. The incident wave on the surface may encounter local facets oriented away from the source in such a way that the local incident angle is below the nominal critical angle. On the contrary, below the nominal critical angle, a large number of local facets oriented toward the source are locally above critical, allowing refraction and therefore enhancing penetration. This effect is illustrated in Fig. 11, which displays the percentage of local facets that are locally above critical versus the nominal grazing angle. Since this calculation is made for a source located at a finite distance from the interface, the effect of the bounded beam is also included. For a flat surface, the bounded beam effect (i.e., the plane wave spectrum is spread around the nominal grazing angle) is the only cause of the nonzero percentage of facets below critical. For nonzero roughness, the percentage of facets above critical remain significant, even at a low grazing angle. This suggests

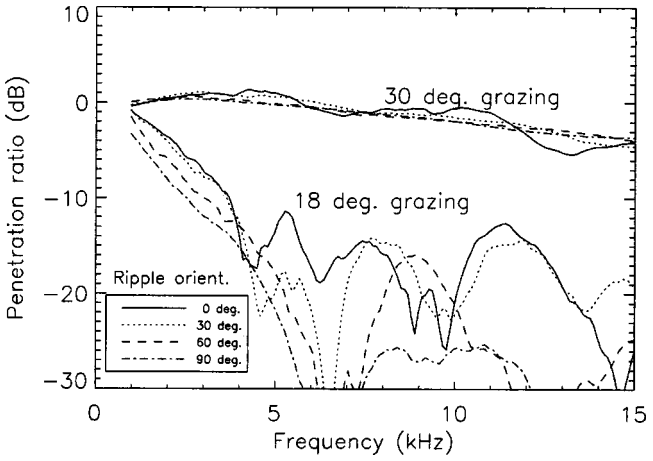


FIG. 14. Example of the variability of the penetration ratio PR(f) for various orientations of the same ripple field (0° corresponds to an insonification across the ripple edges, 90° is along) above critical (top set of curves) and below critical (bottom set of curves). The geometric and geoacoustic parameters are the same as in the previous figures.

that below the critical angle part of the penetration due to roughness is still caused by local refraction. These local refraction effects are more noticeable at 60 cm, where the evanescent wave component becomes weaker. In Fig. 10, at 15°, there is a 6 dB increase in intensity at 30 cm depth and above a 12 dB increase at 60 cm depth. This anomalous penetration mainly due to local refraction is important in terms of the total intensity level, especially in the lower-frequency region. Because refraction seems to play a role well below the nominal critical angle by enhancing penetration, this effect might prove useful for the detection of buried objects.

Figure 12 presents an example of the sensitivity of the penetration ratio to the RMS roughness at various grazing angles. The solid lines correspond to a nonrough seafloor and thus are the results of the combination of the evanescent component with the plane wave spectrum effect.[5] For nonzero roughnesses, it is interesting to observe that as the RMS roughness increases, the penetration ratio increases at lower frequency. This shows that the scattered field and the evanescent wave are of the same order of magnitude at lower frequency for this particular geometry and beam aperture. At higher frequency (i.e., beyond the frequency for which the penetration ratio for zero roughness becomes less than -20 dB), interference fringes keep the penetration ratio to a significant level, but there is very little change in terms of average intensity between 1 cm RMS height and 2 cm RMS height. Also, simulations not displayed here show that the seafloor spatial statistics (mainly the correlation length and the ripple orientation) greatly affect the spectral signature of the received signal by modifying the interference pattern and the amplitude of the penetration ratio. However, the TPR is not significantly affected for various horizontal configurations.

D. Sensitivity of the penetration ratio to seafloor realizations

As suggested in Sec. III B, it is not only the position of the hydrophone but also the seafloor height field realization that greatly affects the scattered component of the received pressure field into the seafloor. Figure 13 shows that each seafloor realization having the same exact statistical characteristics produce a different interference pattern and thus a very different penetration ratio. This suggests also that the scattering field is essentially incoherent as opposed to the evanescent field and nonuniformly distributed in space. In the high-frequency portion of the considered frequency band, the standard deviation of the penetration ratio can be large (about ± 10 dB) but the variation remains much lower in the low-frequency part of the considered frequency band.

E. Influence of roughness anisotropy

Intuitively, a variation of the incoherent field as the ripple orientation changes is expected. Figure 14 shows that insonification of the seafloor across the ripple field produces more interference patterns and higher variance of the scattered field around the coherent field. Along the ripples (90°), the scattered field is very low as if the ripple field was not ''seen'' by the incident sound waves. Between these two extreme cases, there is a gradual change in the level of the scattered field. It is interesting to notice that over the considered frequency band the total penetration ratio (not displayed here) is quasiconstant as the ripple orientation changes above critical within one dB. However, below critical, there is a 5 dB excess of incoherent intensity that penetrates the seafloor across the ripples (0°) compared to the case along the ripples (90°). therefore the aspect dependence of sound penetration into a rough seafloor must be considered for buried object detection.

IV. SUMMARY

A Helmholtz–Kirchhoff (H–K) model of the penetration of sound into sandy seafloors has been presented. With the model we address, in particular, the effects of seabed roughness with particular attention being given to simulation of the two-dimensional seafloor spectrum. The model is restricted by the usual KA assumptions (e.g., no sharp edges, large radii of curvature compared to the wavelength, etc.), but is assumed to be valid over the frequency range and the parameters relevant in the experiment by Maguer on rippled seabeds. The model demonstrates that, as the acoustic frequency increases, the importance of the evanescent wave decreases with concomitant increased importance of roughness scattering and more precisely to local refraction. This latter stands as a viable candidate to explain sound penetration into the seabed at grazing angles beyond ''critical'' for seafloors with ripples. As such studies show good agreement with experimental data, the model may be valid for various ripple heights and orientations.

ACKNOWLEDGMENTS

The authors wish to thank Alain Maguer, Edoardo Bovio, Warren Fox, John Osler, Henrik Schmidt, and Finn Jensen for their helpful comments in the discussions about acoustic seabed penetration.

[1] A. Maguer, E. Bovio, W. L. J. Fox, E. Pouliquen, and H. Schmidt, ''Mechanisms for subcritical penetration into a sandy bottom: experimental and modeling results,'' J. Acoust. Soc. Am. **107**, 1215–1225 (2000).
[2] N. P. Chotiros, ''Biot model of sound propagation in water-saturated sand,'' J. Acoust. Soc. Am. **97**, 199–214 (1995).
[3] N. P. Chotiros, A. M. Mautner, A. Lovik, A. Kristensen, and O. Bergem, ''Acoustic penetration of a silty sand sediment in the 1–10 kHz band,'' IEEE J. Ocean Eng. **22**, 602–615 (1997).
[4] J. L. Lopes, ''Observation of anomalous acoustic penetration into sediment at shallow grazing angles,'' J. Acoust. Soc. Am. **99**, 2473–2474 (1996).
[5] F. B. Jensen and H. Schmidt, ''Subcritical penetration of narrow Gaussian beams into sediment,'' J. Acoust. Soc. Am. **82**, 574–759 (1987).
[6] K. L. Williams, L. J. Satkowiak, and D. R. Bugler, ''Linear and parametric array transmission across a water–sand interface—Theory, experiment, and observation of beam displacement,'' J. Acoust. Soc. Am. **86**, 311–325 (1989).
[7] E. I. Thorsos, D. R. Jackson, J. E. Moe, and K. L. Williams, ''Modeling of subcritical penetration into sediment due to interface roughness, in *High Frequency Acoustics in Shallow Water*, edited by N. G. Pace, E. Pouliquen, O. Bergem, and A. P. Lyons, SACLANTCEN CP-45, Lerici, July 1997.
[8] H. J. Simpson and B. H. Houston, ''A synthetic array measurement of a fast compressional and a slower wave in an unconsolidated water-saturated porous medium,'' J. Acoust. Soc. Am. **102**, 3210 (1997).
[9] D. R. Jackson, K. B. Briggs, K. L. Williams, and M. D. Richardson,

''Tests of models for high-frequency seafloor scattering,'' IEEE J. Ocean Eng. **21**, 458–470 (1996).
[10] J. F. A. Sleath, *Sea Bed Mechanics* (Wiley, New York, 1984).
[11] T. Akal and J. Hovem, ''Two dimensional space series analysis for seafloor roughness,'' Mar. Technol. Soc. J. **3**, 171–182 (1978).
[12] K. B. Briggs, ''Microtopographical roughness of shallow-water continental shelves,'' IEEE J. Ocean Eng. **14**, 360–367 (1989).
[13] R. A. Wheatcroft, ''Temporal variation in bed configuration and one-dimensional bottom roughness at the mid-shell STRESS site,'' Continental Shelf Res. **14**, 1167–1190 (1994).
[14] K. B. Briggs, ''High-frequency acoustic scattering from sediment interface roughness and volume inhomogeneities,'' Naval Research Laboratory, Stennis Space Center, MS, Tech. Report No. NRL/FR/7431-94-9617, 1994.
[15] D. R. Jackson and K. B. Briggs, ''High-frequency bottom backscattering: Roughness versus sediment volume scattering,'' J. Acoust. Soc. Am. **92**, 962–977 (1988).
[16] T. Yang and S. L. Broschat, ''A comparison of scattering model results for two-dimensional randomly rough surfaces,'' IEEE J. Ocean Eng. **40**, 1505–1512 (1992).
[17] M. F. Barnsley, R. L. Devaney, B. B. Mandelbrot, H. O. Peitgen, D. Saupe, and R. F. Voss, *The Science of Fractal Images* (Springer-Verlag, Berlin, 1988).
[18] C. S. Clay and H. Medwin, *Acoustical Oceanography: Principles and Applications* (Wiley, New York, 1970), Chap. 10, Appendix 10.
[19] E. Pouliquen, O. Bergem, G. Canepa, and N. G. Pace, ''Time evolution modeling of seafloor scatter, Part 1: Concept,'' J. Acoust. Soc. Am. **105**, 3136–3141 (1999).

Matched-impulse-response processing for shallow-water localization and geoacoustic inversion

Zoi-Heleni Michalopoulou
Department of Mathematical Sciences and Center for Applied Mathematics and Statistics, New Jersey Institute of Technology, Newark, New Jersey 07102

(Received 10 June 1999; revised 9 March 2000; accepted 5 June 2000)

In this paper, impulse response matching is proposed for source localization and environmental inversion. The ocean impulse response is estimated using a cross-correlation procedure applied to data from the propagation of a broadband pulse in a shallow-water environment. Source localization and geoacoustic parameter estimation are then performed through time-domain correlations between the estimated impulse responses at spatially separated phones and synthetic replica impulse responses. The method is both spatially and temporally coherent. Parameter space search uses a hierarchical scheme designed to exploit the sensitivity of the acoustic field to the unknown parameters. Tested on the SWellEX-96 and synthetic data, the proposed method is shown to be more robust than conventional (linear), incoherent, broadband matched field processing. © *2000 Acoustical Society of America.* [S0001-4966(00)02710-7]

PACS numbers: 43.30.Pc, 43.30.Vh, 43.60.Gk [DLB]

I. INTRODUCTION

This paper proposes geoacoustic inversion and source localization using estimates of the impulse response of the ocean. Estimated ocean impulse response functions have previously been used for checking and calibrating estimation processes related to underwater sound propagation[1,2] but have not been used extensively as the tool of choice for inversion.

Geoacoustic properties of the ocean can be estimated through the inversion of acoustic field measurements at one or more hydrophones employing forward sound propagation modeling and principles of signal processing. Matched field processing (mfp) is one of the popular approaches employed in inversion.[3] Conventional mfp evaluates the correlation between received signals and replica fields and identifies the values of the unknown parameters that yield the maximum replica-signal correlation. Originally developed for narrow-band application, there have been extensive and very successful broadband mfp implementations.[4–8]

Realistic source localization applications of mfp are bedeviled by uncertainties: the source spectrum is often unknown with the sound originating from an uncontrolled source; the ocean channel through which the transmitted signal propagates is not accurately known; ranges rather than exact values are available for environmental parameters such as the sound speed profile, the depth of the water column, and the properties of the bottom sediments. Because of these uncertainties, searches are not only performed for the source location but for the environmental parameters as well, since their impact on the field, and consequently on the source location estimates, is substantial. Searches can be exhaustive or can be combined with global optimization techniques for a more efficient (but, occasionally, with a suboptimal, local maximum result) exploration of the parameter space.[8–15]

In this paper, for the inversion for the source location and geoacoustic parameters we use a modified grid search, an efficient hierarchical search scheme that first inverts for the most important parameters of the acoustic field model and then refines the inversion process by searching over the secondary parameters.

An objective function frequently used in mfp is the linear (Bartlett) processor. This estimator, calculating an inner product between real and replica fields in the frequency domain, is very popular because of its simplicity and its robustness with respect to mismatch between assumed and real parameters. Initially used for narrow-band mfp, linear mfp has also been implemented in many broadband variants by computing arithmetic or geometric means of narrow-band linear ambiguity surfaces. For data received at L hydrophones, the broadband incoherent ambiguity surface $\mathbf{P}_{\text{lin}}$ can be expressed as follows:

$$\mathbf{P}_{\text{lin}}(\mathbf{q}) = \sum_{f=1}^{F} \mathbf{H}_f(\mathbf{q})^* \mathbf{R}_f \mathbf{R}_f^* \mathbf{H}_f(\mathbf{q}), \tag{1}$$

where F is the total number of frequencies considered, $\mathbf{H}_f$ is the L-dimensional vector of the transfer function of the ocean at frequency f, $\mathbf{R}_f$ is the vector of L real data observations at frequency f, $\mathbf{q}$ is the vector of uncertain parameters for which we invert, and $[\cdot]^*$ stands for conjugate transpose.

This paper discusses geoacoustic inversion using the broadband linear estimator and an efficient time-domain inversion technique based on matching an estimate of the real ocean impulse response and replica impulse responses as shown in Refs. 16–19. The latter method follows on the steps of time-series matching for inversion,[1,20–26] theoretically preferable to conventional incoherent broadband mfp since it includes information on the source spectrum and the temporal structure of the acoustic field. In practice, time-series matching is a very computationally demanding process and is therefore rarely used. Impulse response matching, which is the proposed method of this paper, has the advantage of time-series matching in that it also exploits temporal coherence, but is more computationally efficient since the ocean impulse response is usually significantly shorter than

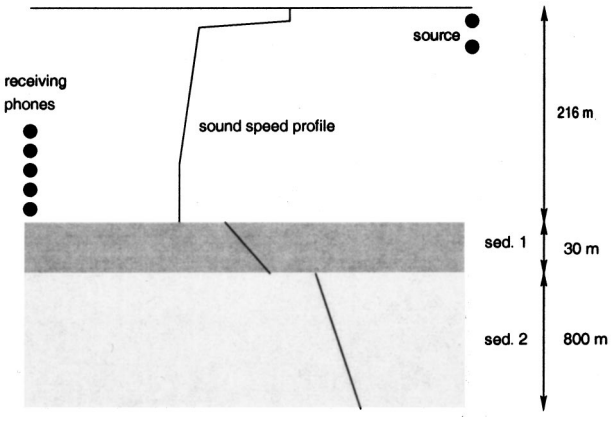


FIG. 1. A sketch of the SWellEX-96 environment.

the long (large time-bandwidth) time series that are transmitted and received in many sonar applications. The method is motivated by the SWellEX-96 data: including linear-frequency modulated (lfm) pulse transmissions, such data are ideal for demonstrating the time-series compression scheme.

The paper is organized as follows: Section II describes the SWellEX-96 data and environment. Section III discusses model-based matched filters, from which the matched-impulse-response processor is derived. Section IV presents the impulse response extraction scheme. Section V describes the inversion approach and contains the parameter estimates obtained with the SWellEX-96 data for near-range data. Section VI discusses results with longer range data. Section VII discusses simulation results. Conclusions follow in Sec. VIII.

II. THE SWellEX-96 DATA

Geoacoustic inversion is performed with data from the SWellEX-96 experiment corresponding to transmissions of lfm signals with broadband information between 200 and 400 Hz. The SWellEX-96 site has been studied extensively and is known quite well.[27–29] A nominal description of the environment incorporates a water column slightly deeper than 200 m, a thin first sediment, a thicker (approximately 800 m) second sediment, and a subspace below. Inversion is performed with the data of event S5. For data collected for a 1-km distance between source and receivers, the water depth did not vary significantly. Thus range-independent modeling is admitted without introducing significant error. The SWellEX-96 environment used as a baseline for the inversion is shown in Fig. 1. Table I contains the ground truth values for the source location coordinates and environmental parameters of the site.

Data recorded at one of 64 vertically separated hydrophones is shown in Fig. 2. One lfm pulse in the 200- to 400-Hz frequency range and nine tones between 109 and 385 Hz are indicated in the figure. The tones are transmitted from a shallow-water source (around 9 m in depth) towed along with the source of interest (at approximately 55 m in depth). The shallow source is here seen as interference. Inversion is performed using the deep source information and 21 out of the 64 phones.

TABLE I. Source location and environmental parameters for SWellEX-96 near-range data.

Source range (km)	1.1
Bottom depth (m)	216.5
Source depth (m)	54.6
Thickness of sed. 1 (m)	23.5
Compressional speed at top of sed. 1 (m/s)	1572
Compressional speed at bottom of sed. 1 (m/s)	1593
Compressional attenuation in sed. 1 (dB/m kHz)	0.2
Density of sed. 1 (g/cm^3)	1.76
Thickness of sed. 2 (m)	800
Compressional speed at top of sed. 2 (m/s)	1881
Compressional speed at bottom of sed. 2 (m/s)	3245
Compressional attenuation in sed. 2 (dB/m kHz)	0.06
Density of sed. 2 (g/cm^3)	2.06

III. INVERSION WITH MODEL-BASED MATCHED FILTERS

A signal transmitted in the ocean interacts with the underwater propagation medium before it is recorded at a receiving phone. The received signal is, therefore, the convolution of the transmitted sequence and the impulse response of the ocean between source and receiver with an additional noise/interference component, functionally expressed as

$$r(t)=\int_{-\infty}^{\infty} s(\tau)h(t-\tau)d\tau+n(t), \tag{2}$$

where $r(t)$ and $s(t)$ are the received and transmitted signals, $h(t)$ is the ocean impulse response, and $n(t)$ is additive noise.

In active problems, the source signature is known. Assuming that the noise is white and Gaussian, optimal localization and inversion for unknown environmental parameters at the receiving sensors is achieved by matched filtering, correlating the received signal to the signal that would be theoretically received, namely the convolution of the source signature and the ocean impulse response.

Obtaining an accurate match requires a process similar to mfp, termed the model-based matched filter (mbmf).[1,23] Specifically, model-based matched filtering requires the generation of replicas of the ocean impulse response for many candidate values of the unknown parameters and for several

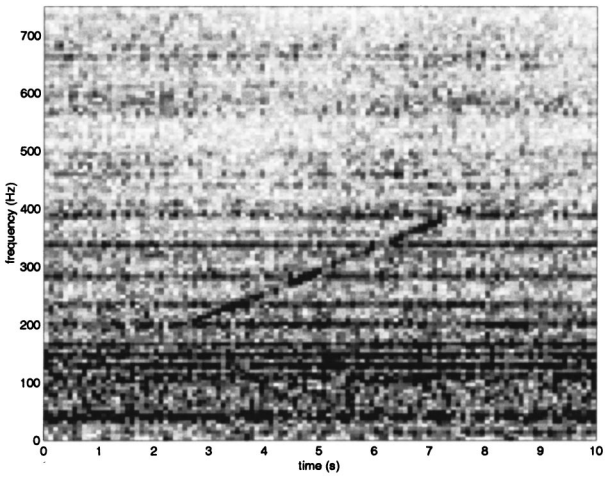


FIG. 2. A pulse received from an lfm transmission.

spatially separated hydrophones; each replica is then convolved with the source signature. The results of the convolution are the received signal replicas, which are cross-correlated to the actual received signals. Cross-correlations are averaged across different phones, and maximizing the resulting average cross-correlation yields estimates of the unknown parameters.

Model-based matched filtering is very similar to matched-field processing in the sense that both approaches rely on the matching of actual to theoretically predicted fields. There is, however, an important difference between the two methods: the mbmf is both temporally and spatially coherent (spatial coherence is present if cross correlations over several spatially separated hydrophones are aggregated), while conventional incoherent broadband mfp is spatially coherent only. It has been previously shown that the exploitation of the additional coherence factor in matched-field approaches leads to improved inversion results.[5,6,17,30–32] Also, the mbmf has been shown to be the optimal estimator of source location and environmental parameters in the case of a known source signature.[23]

In the SWellEX-96 case, the source signature is known; the transmitted signals are lfm pulses carrying information between 200 and 400 Hz. The sampling rate is 1500 Hz and the transmitted pulses are 5-s long. With these parameters, modeling the transmitted lfm signals requires 7501 samples. The received time series will be of a longer duration than the transmission, since the impulse response has a spreading effect on the transmitted signal. For environments similar to that of the SWellEX-96 experiment, the impulse response of the ocean has a duration of approximately 301 samples, generating a received time series of 7801 samples. If inversion were performed by matching the complete received time series to replica time series, cross-correlations between long sequences would be necessary. The process is feasible in the frequency domain, but even there implementation is computationally intensive. Frequency-domain processing requires a Fourier transform of the time-domain data to the frequency domain. Because the time-domain signals consist of many samples, the fields in the frequency domain will be calculated for a large number of frequencies that are very finely spaced. We would then need to generate many replicas (corresponding to the different frequencies) in order to match the received, frequency-domain acoustic fields. Alternatively, frequency-domain processing could be achieved with a coarser frequency spacing adequate for the computation of the transfer function followed by (i) an inverse Fourier transform for the impulse response computation, (ii) convolution between the impulse response and the transmitted time series, and (iii) a Fourier transform for the calculation of the field in the frequency domain. In short, space-time coherent matched filtering carries a computational overhead regardless of the domain in which it is implemented. The expected improvement in performance comes at the cost of expensive implementation.

IV. IMPULSE RESPONSE ESTIMATION

The autocorrelation function of an lfm sequence with frequency content between 200 and 400 Hz is shown in Fig. 3; it can be found to be a sinc pulse approximating closely a δ function.[33] The similarity between the lfm autocorrelation and an impulse can be assessed quantitatively by calculating the effective bandwidth of an lfm sequence; for the lfm sequence considered here, the effective bandwidth is 200. The effective bandwidth, thus, covers the whole range of frequencies (between 200 and 400 Hz) present in the lfm pulse, indicating the strong similarity of the autocorrelation and an impulse.[33,34] For this reason, if we cross-correlate the received time series with the transmitted sequence (the lfm pulse), the result will be the convolution of an approximate δ function and the impulse response of the medium,

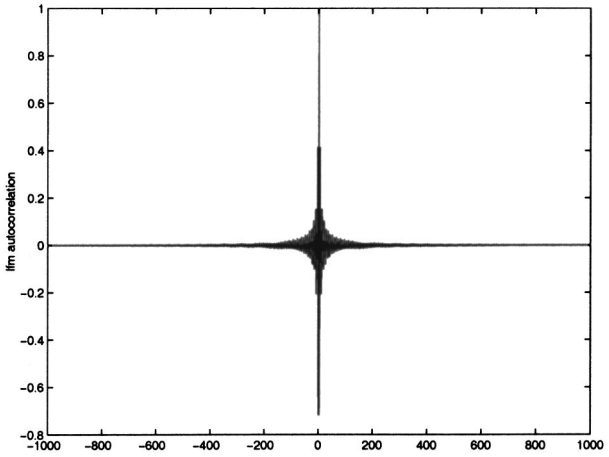


FIG. 3. The autocorrelation function of an lfm pulse.

$$
\begin{aligned}
\hat{h}(t) &= \int_{-\infty}^{\infty} r(\tau)s(\tau-t)d\tau \\
&= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} s(\tau-\alpha)h(\alpha)d\alpha\right]s(\tau-t)d\tau \\
&\quad + \int_{-\infty}^{\infty} n(\tau)s(\tau-t)d\tau \\
&= \int_{-\infty}^{\infty} h(\alpha)\left[\int_{-\infty}^{\infty} s(\tau-\alpha)s(\tau-t)d\tau\right]d\alpha \\
&\quad + \int_{-\infty}^{\infty} n(\tau)s(\tau-t)d\tau \\
&\approx \int_{-\infty}^{\infty} h(\alpha)\,\delta(t-\alpha)d\alpha = h(t), \qquad (3)
\end{aligned}
$$

where $\hat{h}(t)$ is the estimated ocean impulse response and $r(t)$ and $s(t)$ are the received and transmitted sequences. Integral $\int_{-\infty}^{\infty} n(\tau)s(\tau-t)d\tau$ is small compared to the other term in the sum for a high signal-to-noise ratio.

Figure 4 shows a synthetic impulse response of the ocean environment shown in Fig. 1, the estimated impulse response for the same environment obtained with the method described by Eq. (3), and the cross correlation between the original and estimated responses. (The values of the source location and environmental parameters employed in the simulations are summarized in Table I.) The correlation

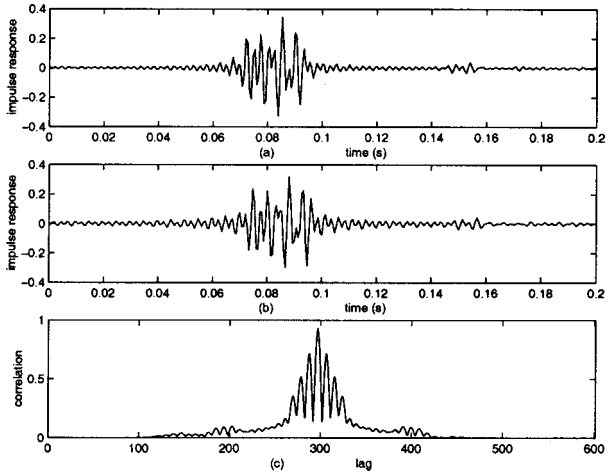


FIG. 4. (a) Synthetic impulse response using normal modes and Fourier synthesis. (b) Estimated impulse response using Eq. (3). (c) Cross correlation between the two impulse responses.

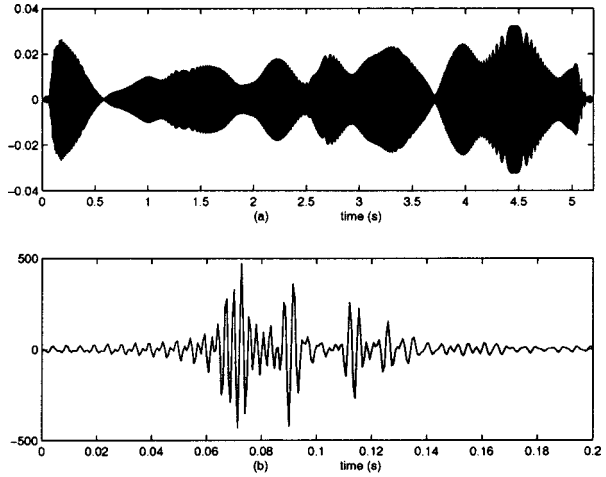


FIG. 6. (a) The signal received at a hydrophone after an lfm transmission and (b) the estimated impulse response.

reaches a level of 95%. The 5% loss occurs because the lfm autocorrelation function is not *identical* to an impulse, although it is very similar.

Impulse response estimation is feasible with the SWellEX-96 data because the signal-to-noise ratio is high and identification of the impulse response sequence is possible from the output of Eq. (3) (that is, the noise term is small). Impulse response extraction can also be performed using deconvolution of the transmitted signal from the received signal. The impulse responses obtained with deconvolution and cross-correlation from the SWellEX-96 data were found to be almost identical. Deconvolution is a more complicated process, however, and the estimation through the cross-correlation calculation is preferred here. More details on deconvolution versus cross-correlation for impulse response estimation can be found in Ref. 34.

Figure 5 shows a synthetic impulse response generated with normal modes[35] and Fourier synthesis for a nominal environment for the site and an estimated impulse response extracted from the received time series at a single hydrophone from the SWellEX-96 data. Although the impulse responses are not identical, they are very similar; cross-correlating the two sequences yields a correlation of 0.73. The two responses represent the replica time series and the data time series that we are trying to best match using the matched-impulse-response processor of this paper. The ambiguity surface, $\mathbf{P}_{\text{imp}}$, matching the two sequences can be written as

$$\mathbf{P}_{\text{imp}}(\mathbf{q}) = \max_{\tau} \left| \sum_{l=1}^{L} \int_{t} \hat{h}_l(t) h_l(t-\tau;\mathbf{q})\, dt \right|, \tag{4}$$

where $\hat{h}_l(t)$ is the estimated impulse response for hydrophone l and $h_l(t)$ is the replica impulse response for the same hydrophone. As in Eq. (1), $\mathbf{q}$ is the vector of parameters for which we invert.

Figure 6 shows the received time series after a transmission of an lfm pulse and the extracted impulse response using Eq. (3). The data signal spans 5.2 s whereas the impulse response has a duration of only 0.2 s, demonstrating that the impulse response estimation process acts as a data compression scheme.

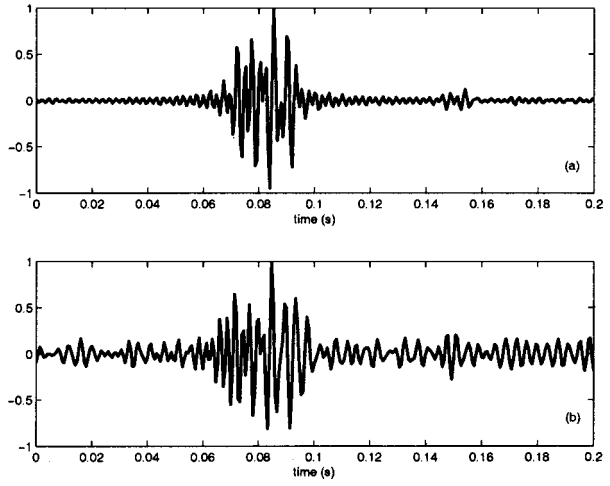


FIG. 5. Comparison of (a) a synthetic impulse response and (b) an estimated impulse response from the SWellEX-96 data.

V. PARAMETER ESTIMATION

Geoacoustic and source related parameters are estimated by identifying which parameter values give the best match between the received signal and computed field replicas. The number of unknown parameters in such a problem can be overwhelmingly large. All parameters, however, are not equally important in their contribution to the acoustic field.[36] The sensitivity of the field to each parameter varies and could depend on the specific structure of the problem. Using this sensitivity to express the relative importance of the parameters to the acoustic field, one can decrease the search space size and implement an efficient inversion approach.[9]

In this work, we identify a hierarchy of the unknown parameters based on the measure of their importance to the acoustic field and perform a layered search for the parameter values that lead to the best match between data and replicas. The first layer consists of a search on a coarse grid among source range and depth, water column depth, first sediment

thickness, compressional speed and attenuation, and density, and corresponding parameters in the second sediment. Subsequently, a search on a fine grid is performed for the parameters that appear to have the most impact to the acoustic field, while the less important parameters were fixed. Once estimates for the ''significant'' (as identified from the first layer of the search) parameters are obtained, a search is performed for values of the secondary set of unknown parameters.

The hierarchical scheme was validated with a simulation study of the behavior of the acoustic field in an environment similar to that of SWellEX-96 subject to variations of the parameters. Specifically, through simulations, we studied acoustic field changes as a function of changes in parameters such as source range and depth, water column depth, and properties of the seafloor sediments.

The above approaches indicated that the three most important parameters in the inversion were the source range, source depth, and bottom depth.

Inversion with both processors involved replica calculation at 41 frequencies between 200 and 400 Hz with a 5-Hz spacing and at 21 phones. The replicas were calculated with KRAKEN.[35] The replica impulse responses were synthesized through an inverse Fourier transform of the frequency domain fields obtained with KRAKEN.

Figure 7 shows the source range and depth ambiguity surfaces obtained with the two processors of Eqs. (1) and (4). Boxes indicate the peak locations on the surfaces. Figure 7(a) shows the source range-depth ambiguity surface computed with the linear processor; the main lobe corresponds to 1.12 km and 60 m in range and depth, respectively, as also indicated in Table II. These source location values are similar to those obtained with the matched-impulse-response processing (listed in the same table). The incoherent ambiguity surface demonstrates a higher level of uncertainty, however, in comparison to the matched-impulse-response ambiguity surface shown in Fig. 7(b). A very strong sidelobe in the surface of Fig. 7(a) can be observed at 1.48 km in range and 9 m in depth. This sidelobe might be a result of the tonal source interference (the tonal source was located approximately at 9 m in depth, as described in Sec. II, its range was the same as the range of the deep source). Such strong interference is not evident in the matched-impulse-response processing ambiguity surface.

It is interesting to observe the behavior of the source range-bottom depth ambiguity surfaces for a fixed source depth, shown in Fig. 8. Both ambiguity surfaces (corresponding to the two processors) are similar, exhibiting a correlation pattern between the source range and the depth of the water column. Both processors identify well the range of the source whereas the depth of the water column is less well defined.

Table II summarizes the source localization results of the two processors and also contains the water-column depth estimates.

Once water-column depth and source location were estimated, the inversion proceeded with sediment properties. Table III presents inversion results for the first sediment properties. No estimates were obtained for the sediment density, since the field exhibited no sensitivity to this parameter.[37] A comparison between Tables I and III shows deviations between estimates and ground truth values for the sediment parameters.

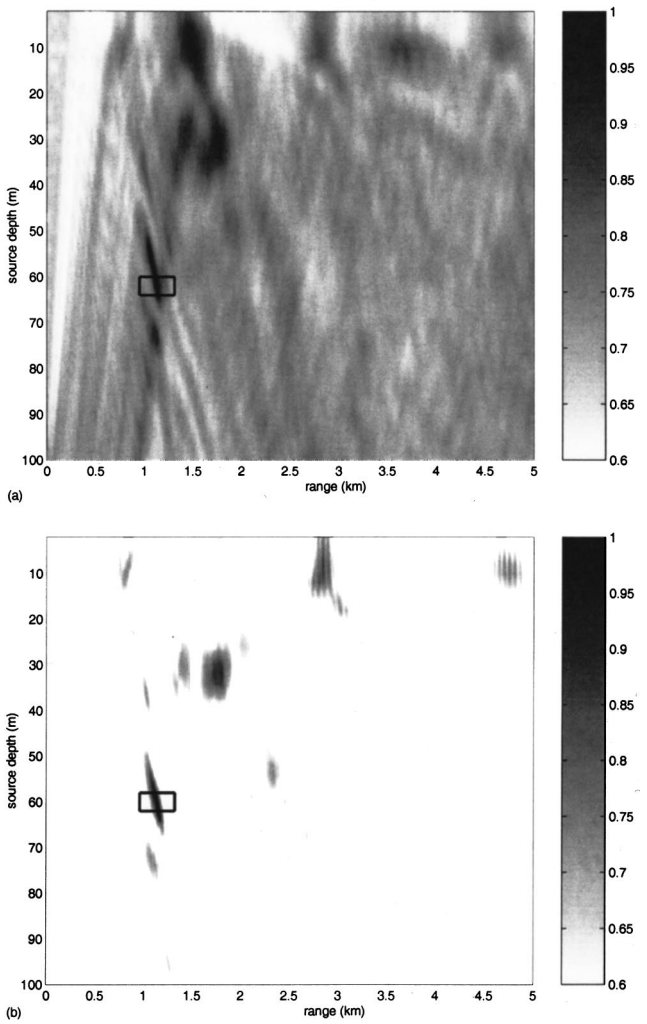


FIG. 7. Source depth-range ambiguity surface obtained with the incoherent linear and the matched-impulse-response processor.

In summary, the source location and water-column depth estimates obtained both with the matched-impulse response and the incoherent processors were consistent with prior knowledge on these parameters. Compressional attenuation for the first sediment was identical for both processors. There were variations in the first sediment thickness, which both processors overestimated in comparison to prior knowledge on the site (attenuation was overestimated as well). The environmental structure found here was, however, similar to

TABLE II. Source range, source depth, and bottom depth estimates obtained with the incoherent linear processor and the matched-impulse-response processor for 1-km data.

	Incoherent linear processor	Matched-impulse-response processor
Source range (km)	1.12	1.06
Source depth (m)	60	56
Bottom depth (m)	214	216

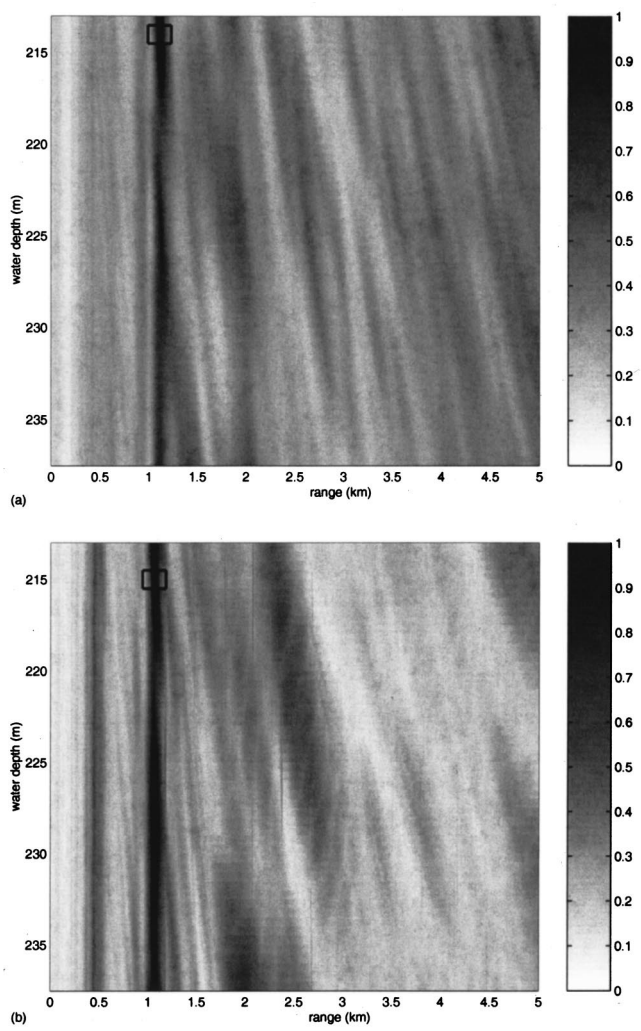


FIG. 8. Water-column depth-source range ambiguity surfaces obtained with (a) the linear and (b) the matched-impulse-response processor. Source depths of 60 and 56 m were used for the linear and matched-impulse response processors, respectively.

the results reported in Ref. 38 for the same data. It is important to note that field sensitivity to the parameters of Table III for the examined 1-km data set was relatively small. Thus the sediment property estimates obtained from this segment of the data might be very vulnerable to noise and might give only a rough description of the sediment.

No results are reported for second sediment properties, because the calculated correlations (for both incoherent linear and matched-impulse-response processors) showed no sensitivity to these parameters.

TABLE III. First sediment parameter inversion with the incoherent linear processor and the matched-impulse-response processor for 1-km data.

	Incoherent linear processor	Matched-impulse-response processor
Attenuation (dB/m kHz)	1.6	1.6
Sediment thickness (m)	100	220
Top compressional speed (m/s)	1560	1601
Bottom compressional speed (m/s)	1609	1601

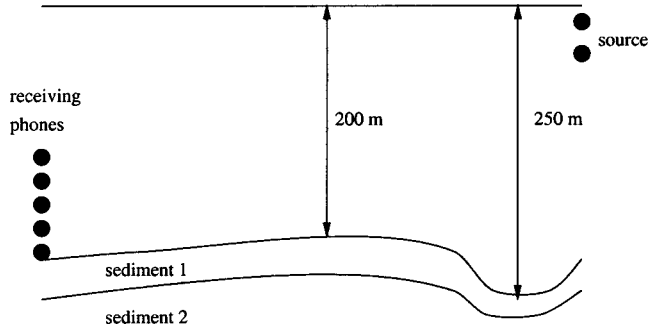


FIG. 9. A sketch of the range-dependent bathymetry for longer range data.

VI. LONG-RANGE RESULTS

Estimation techniques are excellent when they are tested under conditions identical to those under which they were designed, but their performance often deteriorates when deviations of assumptions from actuality are considered. In this section we study the performance of the matched-impulse-response processor and conventional incoherent mfp under mismatch. Inversion results described in Sec. V were obtained from near-range data (the source was approximately 1 km away from the receivers). The inversion was performed with range-independent modeling, a reasonable assumption since the environment did not show dramatic changes over that range. Now we use the environmental parameters (water-column depth and sediment properties) obtained from the near-range data to perform source localization from longer range measurements (close to 8.5 km). The range independence assumption is not an obvious choice here; careful observation of the site map in Ref. 27 shows that steep bathymetric changes occur between source and receiver in the 8.5-km case; the bathymetry of the site is sketched in Fig. 9. The ambiguity surface of Fig. 10(b) shows that the impulse-response matching results are satisfactory, although the environmental changes have not been taken into account. The source location is found to be at a depth of 56 m and a range of 9.22 km. The source depth estimate is very close to the true depth (around 55 m). The source range is shifted, probably because of the mismatch in the bathymetry assumptions (effects of source location shifts in the presence of bathymetry mismatch have been reported in Refs. 39 and 40).

Performance of the conventional incoherent broadband mfp method is much less satisfactory, resulting in a source range-depth ambiguity surface dominated by uncertainty [Fig. 10(a)]. The surface is maximized for a source range and depth of 9.73 km and 32 m, values far from the real source position. The results of both processors for source range and source depth are summarized in Table IV.

Following source localization using the parameters obtained from the near-range data, we performed simultaneous localization and geoacoustic inversion for the 8.5-km data using range-independent modeling and the incoherent linear and matched-impulse-response processors. The localization results were practically identical to those of Table IV. The bottom depth and first sediment results are presented in Table V.

As seen from Table V, the first sediment thickness was found with the matched-impulse-response processor to be 15

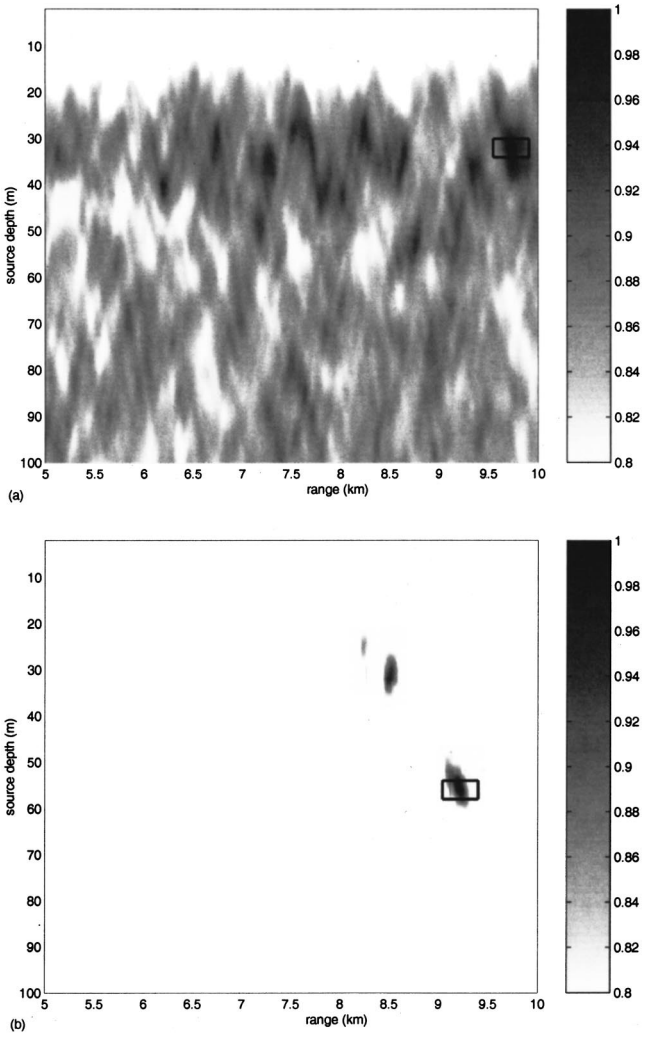


FIG. 10. Ambiguity surfaces from (a) the incoherent linear processor and (b) the impulse response matching process for longer range data.

m and the attenuation coefficient 0.5 dB/m kHz. These values are closer to the known description of the SWellEX-96 site (Fig. 1, Table I, and Refs. 28 and 41) than those of the inversion with the 1-km data. The parameter estimates obtained with the incoherent processor are still quite different from those in Table I and Refs. 28 and 41.

Summarizing the real data results obtained so far, the near-range source localization and bathymetry estimates are very close to the known range and depth of the source emitting the lfm pulses and the known water-column depth, respectively. Very similar results were obtained with both the incoherent linear and the matched-impulse-response processors. In the longer range case, the matched-impulse-response processor gave good source localization results (correct source depth and shift in source range), whereas the incoherent linear processor had a poor localization performance. Deviations were, however, observed between the geoacoustic parameter estimates obtained here (for both near- and far-range data) and the ground truth values of Table I. As mentioned before, one reason for the deviations is the lack of sensitivity of the field at the particular site to sediment parameters.

It should be finally recalled that the field measurements used for the inversion in this work contain interference from a tonal source emitting in parallel to the source of interest, as discussed in Sec. II. This interference is likely to lead to errors in the inversion.

TABLE IV. Source range and source depth estimates obtained with the incoherent linear processor and the matched-impulse-response processor for 8.5-km data. The environmental parameters were set to values obtained from the near-range data inversion.

	Incoherent linear processor	Matched-impulse-response processor
Source range (km)	9.73	9.22
Source depth (m)	32	56

TABLE V. Water column depth and first sediment parameter estimates obtained with the incoherent linear processor and the matched-impulse-response processor for 8.5-km data.

	Incoherent linear processor	Matched-impulse-response processor
Bottom depth (m)	225.5	216.5
Attenuation (dB/mk Hz)	1.5	0.5
Sediment thickness (m)	135	15
Top compressional speed (m/s)	1570	1537
Bottom compressional speed (m/s)	1630	1630

VII. SIMULATIONS

The results presented in the previous sections suggest superiority of the matched-impulse-response processor over the conventional incoherent linear processor. To further study the comparative performance of the two processors we performed simulation experiments focused on source range localization. The source was located at a range of 1.1 km and a depth of 54 m and emitted sound at frequencies between 200 and 400 Hz; the environment of Fig. 1 and Table I was used for the simulations and 21 receiving phones were considered. The sampling rate was 1500 Hz. The criterion for comparing the two processors was the probability of correct localization, which was here defined as the probability that the estimated range falls within 10% of the true value.

For both processors, replicas were generated at 41 frequencies using the normal modes approach, as described in Sec. V. The frequency-domain replicas generated for the 41 frequencies between 200 and 400 Hz with a 5-Hz spacing allowed the generation of unaliased impulse response replicas used in the matched-impulse-response processing. The linear processor involved matching of the frequency-domain field at the same 41 frequencies to the computed frequency-domain replicas. In this way, both processors required exactly the same number of field computations.

The results of a Monte Carlo simulation were very different for the two processors and perhaps surprising at first glance. For a signal-to-noise ratio of 10 dB, the matched-impulse-response processor had excellent (100% correct) localization, which was expected; the probability of correct localization, however, for the incoherent processor was only 22%. A careful inspection of the two processors explains the substantial difference in performance; the received signal in

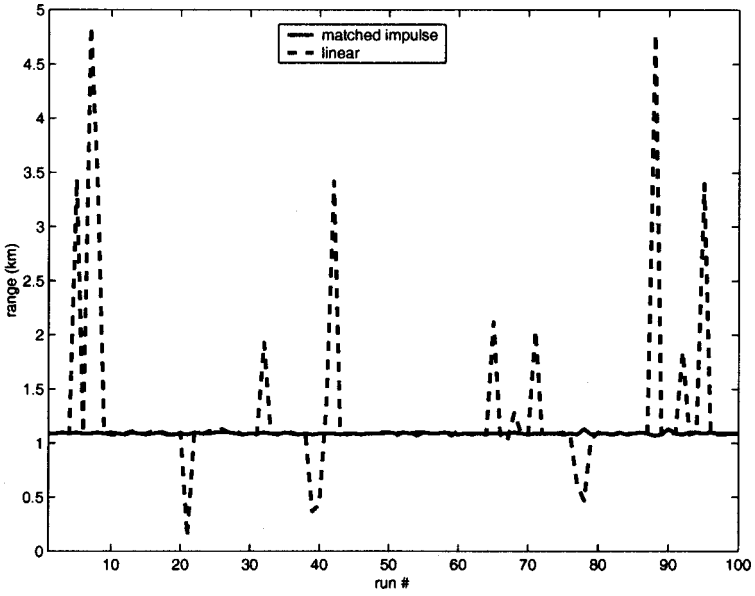


FIG. 11. Synthetic source range estimation results obtained from 100 data realizations using the matched-impulse response and linear processors.

a case simulating the SWellEX-96 experiment consists of 7801 samples (for a 301 long impulse response) as discussed in Sec. III; the Fourier transform of the data should also consist of at least the same number of samples. For a signal with content between 200 and 400 Hz, the received energy would then be distributed among 1041 frequencies with a very fine spacing between them. For the linear processor we only keep 41 out of these 1041 frequencies for the matching process, reducing dramatically the signal-to-noise ratio, since we ignore many samples carrying signal information (this information loss took place in the real data processing as well). In order to account for the signal information that we now ignore in the processing, we would have to generate replicas at 1041 frequencies instead of 41. The process would then be very computationally intensive. No such information loss occurs in the matched-impulse-response processing scheme. The signal energy is mapped from the received signal to the impulse response [Eq. (3)] which is then matched to replica impulse responses.

In order to further compare the two processors in the absence of signal-to-noise ratio inequalities, we performed additional simulations. Both processors used replicas at the same 41 frequencies as before, only now no signal information was neglected for the linear processor because a short duration signal was considered at the receiver. Specifically, the received signal this time was much shorter than the received signal in the SWellEX-96 experiment and the previous simulations; it consisted of 301 samples, leading to 301 point Fourier transforms. For a 301 sample frequency domain representation for a signal with a 1500-Hz sampling rate and content between 200 and 400 Hz, the signal energy is contained in 41 frequency samples between 200 and 400 Hz. For a signal-to-noise ratio of 4 dB, a probability of correct localization of 1 was calculated for the matched-impulse-response processor, whereas a value of only 0.85 was obtained for the linear processor; the results of 100 runs are shown in Fig. 11. More extensive simulations pointed to the same general conclusion that the impulse response processor is a superior estimator to the conventional, incoherent processor even when the processors use exactly the same amount of signal information. The performance gap between the two processors increased, as longer duration signals were considered at the receivers and more signal information was neglected in the incoherent linear processor implementation.

VIII. CONCLUSIONS

Matched-field inversion was performed using an efficient, temporally and spatially coherent technique. The inversion approach of this paper is based on estimation of the impulse response of the ocean, which is feasible when the source signature is known and facilitated when the autocorrelation of the source signature approximates an impulse (δ function). The estimated impulse response is then matched to replica impulse responses for parameter estimation. The matched-impulse-response processing technique was combined with a hierarchical inversion scheme, estimating parameters in ''layers'' depending on the magnitude of their effect on the acoustic field. The proposed method gave excellent source localization results and reasonable environmental parameter estimates when applied to SWellEX-96 data. These results were compared with and found superior to conventional, incoherent matched-field processing results. Methods other than the cross-correlation operation are currently being examined with the goal of more accurate extraction of the ocean impulse response.

The matched-impulse-response processing scheme discussed here is both spatially and temporally coherent in contrast to the only spatially coherent linear processor. As has been discussed in relation to coherent processors in matched field processing,[5,6,17,30–32,42] the additional coherence allows the processor to exploit more information. In addition, using exactly the same replicas as the incoherent linear processor, the matched-impulse-response processor makes better use of the available signal information than the linear processor; signal energy at a large number of frequencies is not taken into account in the incoherent linear processing, whereas all

the signal energy is used in the matched-impulse-response processing approach.

ACKNOWLEDGMENTS

The author wishes to thank Dr. Newell Booth for providing the acoustic data, Dr. W. Hodgkiss and Dr. James Murray for their invaluable help with the data, and Dr. Paul Baxley for fruitful discussions. This work was supported by ONR through Grant No. N00014-97-1-0600, NSF through Grant No. DMS-9872008, and NJIT through SBR. The SWellEX-96 experiment was supported by ONR.

[1] J.-P. Hermand, ''Broadband geoacoustic inversion in shallow water from waveguide impulse response measurements on a single hydrophone; theory and experimental results,'' IEEE J. Ocean Eng. **24**, 41–66 (1998).
[2] P. Mignerey and S. Finette, ''Multichannel deconvolution of an acoustic transient in an oceanic waveguide,'' J. Acoust. Soc. Am. **92**, 351–364 (1992).
[3] A. Tolstoy, *Matched Field Processing for Underwater Acoustics* (World Scientific, Singapore, 1993).
[4] A. Baggeroer, W. Kuperman, and H. Schmidt, ''Matched field processing: Source localization in correlated noise as an optimum parameters estimation problem,'' J. Acoust. Soc. Am. **83**, 571–587 (1988).
[5] Z.-H. Michalopoulou and M. B. Porter, ''Matched-field processing for broadband source localization,'' IEEE J. Ocean Eng. **21**, No. 4, 384–392 (1996).
[6] E. K. Westwood, ''Broadband matched-field source localization,'' J. Acoust. Soc. Am. **91**, 2777–2789 (1992).
[7] J. A. Shorey and L. W. Nolte, ''Wideband optimal *a posteriori* probability source localization in an uncertain shallow ocean environment,'' J. Acoust. Soc. Am. **103**, 355–361 (1998).
[8] P. Gerstoft and D. Gingras, ''Parameter estimation using multi-frequency range dependent data in shallow water,'' J. Acoust. Soc. Am. **99**, 2839–2850 (1996).
[9] W. A. Kuperman, M. D. Collins, J. S. Perkins, and N. R. Davis, ''Optimal time domain beamforming with simulated annealing including application of *a priori* information,'' J. Acoust. Soc. Am. **88**, 1802–1810 (1990).
[10] M. D. Collins and W. Kuperman, ''Focalization: Environmental focusing and source localization,'' J. Acoust. Soc. Am. **90**, 1410–1422 (1991).
[11] P. Gerstoft, ''Inversion of seismoacoustic data using generic algorithms and *a posteriori* probability distributions,'' J. Acoust. Soc. Am. **95**, 770–782 (1994).
[12] D. Gingras and P. Gerstoft, ''Inversion for geometric and geoacoustic parameters in shallow water: Experimental results,'' J. Acoust. Soc. Am. **97**, 3589–3598 (1995).
[13] J. Shorey, L. Nolte, and J. Krolik, ''Computationally efficient Monte Carlo estimation algorithms for matched field processing in uncertain ocean environments,'' J. Comput. Acoust. **2**, No. 3, 285–314 (1994).
[14] Z.-H. Michalopoulou and M. B. Porter, ''Focalization in the Gulf of Mexico,'' in ICASSP-96, Vol. 6, May 1996, Atlanta, pp. 3086–3089.
[15] M. Musil, N. R. Chapman, and M. Wilmut, ''Range dependent matched-field inversion of SWellEX-96 data using the downhill simplex algorithm,'' J. Acoust. Soc. Am. **106**, 3270–3281 (1999).
[16] Z.-H. Michalopoulou, M. B. Porter, and J. P. Ianniello, ''Broadband source localization in the Gulf of Mexico,'' J. Comput. Acoust. **4**, 361–370 (1996).
[17] Z.-H. Michalopoulou, ''Active target detection in the ocean: Optimizing performance and cost,'' J. Acoust. Soc. Am. **103**, 2855 (1998).
[18] Z.-H. Michalopoulou and M. Picarelli, ''Coherent and incoherent shallow water matched field inversion,'' J. Acoust. Soc. Am. **104**, 1739 (1998).
[19] M. B. Porter, S. M. Jesus, Y. Stephan, E. Coelho, and X. Demoulin, ''Single-phone source tracking in a variable environment,'' in *Proceedings of the Fourth European Conference on Underwater Acoustics*, Vol. 2, pp. 575–580, 1998.
[20] C. S. Clay, ''Optimum time-domain signal transmission and source location in a waveguide,'' J. Acoust. Soc. Am. **81**, 660–664 (1987).
[21] S. Li and C. S. Clay, ''Optimum time-domain signal transmission and source location in a waveguide: Experiments in an ideal wedge waveguide,'' J. Acoust. Soc. Am. **82**, 1409–1417 (1987).
[22] R. Brienzo and W. S. Hodgkiss, ''Broadband matched-field processing,'' J. Acoust. Soc. Am. **94**, 2821–2831 (1993).
[23] J.-P. Hermand and W. I. Roderick, ''Delay-Doppler resolution performance of large time-bandwidth-product linear FM signals in a multipath ocean environment,'' IEEE J. Ocean Eng. **94**, 1709–1727 (1988).
[24] C.-S. Chiu, J. H. Miller, and J. F. Lynch, ''Forward coupled-mode propagation modeling for coastal acoustic tomography,'' J. Acoust. Soc. Am. **99**, 793–802 (1996).
[25] E. K. Westwood and D. P. Knobles, ''Source track localization via multipath correlation matching,'' J. Acoust. Soc. Am. **102**, 2645–2654 (1997).
[26] M. T. Wazenski and D. Alexandrou, ''Active, wideband detection and localization in an uncertain multipath environment,'' J. Acoust. Soc. Am. **101**, 1961–1970 (1997).
[27] ''SWellEX-96 Preliminary Data Report,'' Technical Report, 1996.
[28] R. T. Bachman, P. W. Schey, N. O. Booth, and R. J. Ryan, ''Geoacoustic databases for matched field processing: Preliminary results in shallow water off San Diego, California,'' J. Acoust. Soc. Am. **99**, 2077–2085 (1996).
[29] N. O. Booth, P. A. Baxley, J. A. Rice, P. W. Schey, W. S. Hodgkiss, G. L. D. Spain, and J. J. Murray, ''Source localization with broadband matched field processing in shallow water,'' IEEE J. Ocean Eng. **21**, 402–412 (1996).
[30] G. J. Orris, M. Nicholas, and J. S. Perkins, ''Coherent and incoherent broadband matched-field processing with SWellEX-96 data,'' J. Acoust. Soc. Am. **102**, 3170 (1997).
[31] Z.-H. Michalopoulou, ''Robust multi-tonal matched-field inversion: A coherent approach,'' J. Acoust. Soc. Am. **104**, 163–170 (1998).
[32] Z.-H. Michalopoulou, ''An efficient implementation of broadband matched field processing for a known source spectrum,'' in *Proceedings of the Fourth European Conference on Underwater Acoustics*, Vol. 1, pp. 45–50, 1998.
[33] W. S. Burdic, *Underwater Acoustic System Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1984).
[34] Z.-H. Michalopoulou, ''On the estimation of the ocean impulse response,'' CAMS Research Report 9900-7, 1999.
[35] M. B. Porter, ''The KRAKEN normal mode program,'' SACLANT Undersea Research Centre Memorandum (SM-245) and Naval Research Laboratory Mem. Rep. 6920, 1991.
[36] J. S. Perkins, M. D. Collins, and L. T. Fialkowski, ''Application of coordinate rotation to inversion for sediment parameters,'' J. Acoust. Soc. Am. **103**, 2936 (1998).
[37] F. Jensen, W. Kuperman, M. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, New York, 1994).
[38] P. A. Baxley, ''Matched field geoacoustic parameter inversion using horizontal line arrays,'' in *Fourth International Conference on Theoretical and Computational Acoustics*, Trieste, Italy, 1999.
[39] D. R. D. Balzo, C. Feuillade, and M. R. Rowe, ''Effects of water depth mismatch on matched-field localization in shallow water,'' J. Acoust. Soc. Am. **83**, 2180–2185 (1988).
[40] G. D'Spain, J. Murray, and W. Hodgkiss, ''Mirages in shallow water matched field processing,'' J. Acoust. Soc. Am. **105**, 3245–3265 (1999).
[41] P. A. Baxley, ''Matched-field bottom property inversion and sensitivity analysis for the swellex series experiments,'' J. Acoust. Soc. Am. **104**, 1740 (1998).
[42] S. P. Czenszak and J. L. Krolik, ''Robust wideband matched-field processing with a short vertical array,'' J. Acoust. Soc. Am. **101**, 749–759 (1997).

Kramers–Kronig relations applied to finite bandwidth data from suspensions of encapsulated microbubbles

Joel Mobley and Kendall R. Waters
Laboratory for Ultrasonics, Department of Physics, Washington University, St. Louis, Missouri 63130-4899

Michael S. Hughes
Mallinckrodt, Inc., 675 McDonnell Boulevard, Hazelwood, Missouri 63042

Christopher S. Hall and Jon N. Marsh
Laboratory for Ultrasonics, Department of Physics, Washington University, St. Louis, Missouri 63130-4899

Gary H. Brandenburger
Mallinckrodt, Inc., 675 McDonnell Boulevard, Hazelwood, Missouri 63042

James G. Miller
Laboratory for Ultrasonics, Department of Physics, Washington University, St. Louis, Missouri 63130-4899

(Received 3 May 1999; accepted for publication 27 July 2000)

In this work, the Kramers–Kronig (K–K) relations are applied to experimental data of resonant nature by limiting the interval of integration to the measurement spectrum. The data are from suspensions of encapsulated microbubbles (Albunex®) and have the characteristics of an ultrasonic notch filter. The goal is to test the consistency of this dispersion and attenuation data with the Kramers–Kronig relations in a strict manner, without any parameters from outside the experimental bandwidth entering in to the calculations. In the course of reaching the goal, the artifacts associated with the truncation of the integrals are identified and it is shown how their impacts on the results can be minimized. The problem is first approached analytically by performing the Kramers–Kronig calculations over a restricted spectral band on a specific Hilbert transform pair (Lorentzian curves). The resulting closed-form solutions illustrate the type of artifacts that can occur due to truncation and also show that accurate results can be achieved. Next, both twice-subtracted and lower-order Kramers–Kronig relations are applied directly to the attenuation and dispersion data from the encapsulated microbubbles. Only parameters from within the experimental attenuation coefficient and phase velocity data sets are used. The twice-subtracted K–K relations produced accurate estimates for both the attenuation coefficient and dispersion across all 12 data sets. Lower-order Kramers–Kronig relations also produced good results over the finite spectrum for most of the data. In 2 of the 12 cases, the twice-subtracted relations tracked the data markedly better than the lower-order predictions. These calculations demonstrate that truncation artifacts do not overwhelm the causal link between the phase velocity and the attenuation coefficient for finite bandwidth calculations. This work provides experimental evidence supporting the validity of the subtracted forms of the acoustic K–K relations between the phase velocity and attenuation coefficient. [S0001-4966(00)01611-8]

PACS numbers: 43.35.Bf [HEB]

I. INTRODUCTION

The derivation of dispersion relations based on causality had its beginnings in 1926 with Kronig's work relating the dispersion and absorption of x rays.[1] In 1927, Kramers showed that the existence of electromagnetic dispersion relations implies that no signal can propagate in a medium faster than the vacuum speed of light (relativistic causality).[2] Eventually, the general causal basis for the dispersion relations was appreciated.[3] By utilizing the physical restrictions on the behavior of stable linear systems, Kramers–Kronig (K–K) relations can be established between the real and imaginary parts of a frequency-domain response function of the system.[3–5] In optics, the K–K transforms are central to the widely used experimental technique of determining the complex refractive index from reflection measurements.[6–8] For acoustic wave propagation, Kramers–Kronig relations between the phase velocity and attenuation coefficient were first derived by Ginzberg.[9] Subsequently, several distinct forms of Kramers–Kronig relations between attenuation and velocity have been proposed and/or applied.[5,10–25] The relations used in this study are derived using Titchmarsh's theorem[26] and the method of subtractions[27] following earlier work.[5,12]

There are a number of publications where K–K relations are applied to data from acoustic/ultrasonic experiments.[10,11,13–17,20,22–25,28–31] In some instances the full relations are employed and either an analytic model is fit to the data or the integrals are restricted to the measurement spectrum. Other studies have used the nearly local form of the relations derived for finite bandwidths.[11,13] Causality-based

methods for linking attenuation and dispersion without explicit use of K–K transforms have also been explored[32,33] (although the same results are derivable from K–K relations as well[10,28,30,31]).

A. Description of the present work

The aim of this work is to test the consistency of the subtracted K–K relations for the complex wave number with data from a resonant system in a strict manner. To that end, the exact K–K integrands are used, but the limits of integration are restricted to the bandwidth of the measurements. Only experimentally determined parameters from the attenuation and dispersion data sets are used in the calculations. The relations are applied directly to the discrete data sets—no fitting to analytical models is performed. Many of the other experimental tests of the local or finite bandwidth K–K relations for ultrasound have involved systems with dispersion and attenuation characteristics that are monotonic functions of frequency (e.g., power law attenuation). In contrast, this work focuses on a resonant system having peaks in both phase velocity and attenuation inside the measurement spectrum. The data are from measurements of dilute Albunex, a pharmaceutical product consisting of a suspension of protein-coated microbubbles. Dilute Albunex suspensions are essentially ultrasonic notch filters where the encapsulated microbubbles behave like damped resonators.

In order to interpret the K–K results, distinguishing between the causal and artifactual influences on the calculations is crucial. Thus, before presenting the outcomes of the microbubble calculations, we investigate the truncation problem theoretically to establish some context for understanding the potential impact of truncation artifacts. From the analytical examination of a specific Hilbert transform pair (the Lorentzian curves) under the twice-subtracted K–K transformations, closed-form expressions for the truncation artifacts are given. Although not an exact match to the behavior of the Albunex data examined later, the Lorentzian pair provides a suitable model for understanding the behavior of resonant systems under the finite-interval K–K transformations. These theoretical results show that good agreement between the predictions and target data can be achieved. They also reveal the potentially dominating artifacts inherent in the twice-subtracted relations and show how these artifacts can be minimized.

B. Related work on K–K and bubbly fluids in acoustics

Recent theoretical/computational studies using acoustic Kramers–Kronig relations include works by Ye[34] and Tempkin[21] involving bubbles in fluid. In Ye's work, the consistency between a multiple scattering dispersion relation derived from physical models and the more general K–K relations was demonstrated. Ye performed the K–K calculations numerically for several model systems of spherical scatterers, including bubbles in liquid. Tempkin used an alternate form of the K–K relations to derive an expression for the phase velocity in a bubbly liquid from an analytical expression for the attenuation. In contrast to our work, neither of these investigations involved the use of experimental data or was concerned with finite restriction artifacts. Our analytical work on the Lorentzian curves are used only to provide a context for discriminating between the causally linked features in the microbubble data and the artifacts of truncation, a different aspect of the problem than addressed in the other papers.

C. Related work on the finite interval problem

There are many studies related to the restriction of K–K relations to finite intervals (e.g., see Refs. 35–38). The study most closely related to our present work involves the phase determination for infrared reflection spectra in optics.[39] This earlier paper provides approximate limits for the artifacts in the phase based on the frequency endpoints and extrema of the reflectance outside their known spectrum. Using an analytical model for inputs to the K–K integrals, they make numerical determinations of the ultimate impact on the calculation of the absorption spectrum with both one and two peaks. No experimental data was used and only approximate analytical results were given, in contrast to the present work.

D. Outline of remaining sections

In Sec. II, we discuss relevant aspects of the Kramers–Kronig theory and its application to the complex wave number for ultrasonic propagation. We also investigate the finite interval problem for K–K calculations. General issues related to the K–K kernel are examined first. We then look at system-specific problems by applying the truncated (i.e., confined to a finite interval of integration) twice-subtracted Kramers–Kronig relations to a model resonant system. We provide closed-form expressions for the artifacts and demonstrate the how the ''linear'' artifact arises and how it can be suppressed. Section III provides a brief discussion of the experimental data and numerical methods used in this work. In Sec. IV the results of applying K–K transformations directly to the experimental Albunex data are presented, using both twice-subtracted and lower-order forms of the K–K relations. In Sec. V, various aspects of the results are discussed. We look at the relative accuracies of the different relations in generating the correct scale and shape of the target experimental data. The central role of the linear artifact in choosing the subtraction frequency is also examined. Some further general issues are briefly addressed, followed by the conclusions. The derivations of the K–K expressions used in this article are provided in Appendix A, while details of the Lorentzian model calculations are given in Appendix B.

II. THEORY

A. General Kramers–Kronig theory

The Kramers–Kronig relations are primarily the consequence of causality—simply that a system cannot react to a stimulus before the stimulus occurs. A system will obey

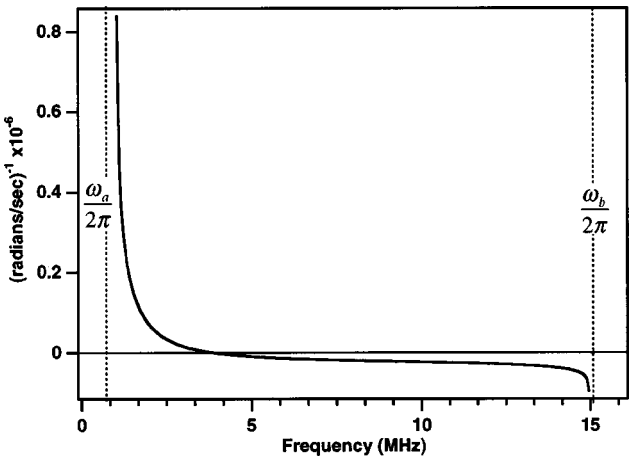


FIG. 1. The integral of the Kramers–Kronig kernel [Eq. (7)] over the finite bandwidth spanned by the Albunex data (1–15 MHz).

K–K relations given that it satisfies the physical requirements of causality, linearity, and finite energy response (i.e., that the energy contained in the system's response to a time-localized stimulus be finite). The response of the system can be represented by a compressibility, a combination of elastic moduli, or some other parametrization appropriate to the problem. The real and imaginary parts of the frequency-domain representation of the system response function (e.g., transfer function) are the quantities related by the K–K transformations. In wave propagation problems, a physically meaningful representation of an extended system is in terms of the complex wave number

$$K(\omega)=\frac{\omega}{c(\omega)}+i\alpha(\omega), \tag{1}$$

where $H(\omega,D)=\exp(iK(\omega)D)$ is the transfer function, $c(\omega)$ is the phase velocity, $\alpha(\omega)$ is the attenuation coefficient, and D is the sample thickness. The complex wave number is proportional to the logarithm of the system response,

$$\frac{\ln H(\omega,D)}{D}=iK(\omega)=-\alpha(\omega)+i\frac{\omega}{c(\omega)}. \tag{2}$$

The relations are established for the components of the complex wave number through the analytic properties inherited from the system response along with the method of subtractions. The details of the derivation are given in Appendix A.

B. Expansion form of the twice-subtracted K–K relations

The relations used here require that subtractions be used to ensure the convergence of the K–K integrals. This is discussed in detail in Appendix A. Here we discuss the ''expansion'' forms of the twice-subtracted K–K relations between $\alpha(\omega)$ and $c(\omega)$ that are used as the starting points for the finite-interval calculations discussed later. In this work, the expansion forms of the relations are of interest primarily because our data are known only over a limited spectrum where they are dominated by localized features (e.g., peaks and nadirs) rather than global trends. By regrouping the terms in Eq. (A10) and transforming the negative frequency portions of the integrals to positive frequencies, we get the twice-subtracted attenuation coefficient relation in the ''expansion'' form,

$$\begin{aligned}\alpha(\omega)=\lim_{|\Omega|\to\infty}\Bigg(&\left[\frac{1}{\pi}\int_0^{\Omega}\frac{\omega'/c(\omega')-\omega/c(\omega)}{\omega'-\omega}d\omega'+\frac{1}{\pi}\int_0^{\Omega}\frac{\omega'/c(\omega')+\omega/c(\omega)}{\omega'+\omega}d\omega'\right]\\&+\left[\alpha(\omega_0)-\frac{1}{\pi}\int_0^{\Omega}\frac{\omega'/c(\omega')-\omega_0/c(\omega_0)}{\omega'-\omega_0}d\omega'-\frac{1}{\pi}\int_0^{\Omega}\frac{\omega'/c(\omega')+\omega_0/c(\omega_0)}{\omega'+\omega_0}d\omega'\right]\\&+(\omega-\omega_0)\left[\frac{d}{d\omega}\alpha(\omega)\Big|_{\omega=\omega_0}-\frac{1}{\pi}\int_0^{\Omega}\mathrm{Im}\,\Lambda_2(\omega',\omega_0)d\omega'-\frac{1}{\pi}\int_0^{\Omega}\mathrm{Im}\,\Lambda_2(-\omega',\omega_0)\,d\omega'\right]\Bigg),\end{aligned} \tag{3}$$

where

$$\mathrm{Im}\,\Lambda_2(\omega,\omega_0)=\frac{\omega/c(\omega)-\omega_0/c(\omega_0)-(\omega-\omega_0)(d/d\omega)\omega/c(\omega)\big|_{\omega=\omega_0}}{(\omega-\omega_0)^2}, \tag{4}$$

and ω_0 is known as the subtraction frequency. [The role of the limiting process in Eq. (3) is discussed in Appendix A.] The first two integrals on the right-hand side together comprise the Hilbert transform which essentially translates the locally varying behavior of one function into local variations of its K–K partner. The singular Hilbert transform kernel $1/(\omega'-\omega)$ strongly weighs the numerator of the first integral when close to ω. Since the kernel is antisymmetric about ω, only the part of the numerator that is odd about ω (e.g., the odd-order terms in the Taylor expansion of the numerator about ω) contributes to the integration. For sufficiently isolated features, this leads to the familiar dispersive structure where the steeply sloped regions in one member correspond to peaks (or nadirs) in its K–K partner. The other terms on the right-hand side assist in accounting for pan-spectral behavior by adding a constant and modifying the variations

with a linear trend. For this work, it is useful to view the terms beyond the initial Hilbert transform as a power series expansion to first order in $\omega-\omega_0$. Going to higher-order relations would add higher-order terms to the series. If the complex wave number obeyed the relations with one subtraction, then the linear term would be identically zero; if it obeyed the Hilbert transforms alone, both expansion terms would be identically zero.

The case of the phase velocity relation is more involved. First, the derivative on the left-hand side of Eq. (A11), $(d/d\omega)\omega/c(\omega)\big|_{\omega=\omega_0}$ is taken before regrouping the terms. The phase velocity relation takes the form

$$\omega\left(\frac{1}{c(\omega)}-\frac{1}{c(\omega_0)}\right)=\lim_{|\Omega|\to\infty}\left(\left[\frac{1}{\pi}\int_0^{\Omega}\frac{\alpha(\omega')-\alpha(\omega)}{\omega'-\omega}d\omega'-\frac{1}{\pi}\int_0^{\Omega}\frac{\alpha(\omega')-\alpha(\omega)}{\omega'+\omega}d\omega'\right]\right.$$
$$-\left[\frac{1}{\pi}\int_0^{\Omega}\frac{\alpha(\omega')-\alpha(\omega_0)}{\omega'-\omega_0}d\omega'-\frac{1}{\pi}\int_0^{\Omega}\frac{\alpha(\omega')-\alpha(\omega_0)}{\omega'+\omega_0}d\omega'\right]$$
$$\left.+(\omega-\omega_0)\left(\omega_0\frac{d}{d\omega}\frac{1}{c(\omega)}\bigg|_{\omega=\omega_0}+\frac{1}{\pi}\int_0^{\Omega}\mathrm{Re}\,\Lambda_2(\omega',\omega_0)\,d\omega'+\frac{1}{\pi}\int_0^{\Omega}\mathrm{Re}\,\Lambda_2(-\omega',\omega_0)\,d\omega'\right)\right),\qquad(5)$$

where

$$\mathrm{Re}\,\Lambda_2(\omega,\omega_0)=\frac{-\alpha(\omega)+\alpha(\omega_0)+(\omega-\omega_0)(d/d\omega)\alpha(\omega)\big|_{\omega=\omega_0}}{(\omega-\omega_0)^2}.\qquad(6)$$

The one difference to note between the attenuation and phase velocity relations involves the linear term. If the linear term in the $\alpha(\omega)$ determination in Eq. (3) is ignored, the remaining terms are equivalent to relation with one subtraction. For the phase velocity [Eq. (5)], this is not the case because we performed the derivative of $\omega/c(\omega)$ before grouping the terms. Later on we refer to a ''lower-order'' relation for the phase velocity which is Eq. (5) with the linear term removed. This point is explored further in Secs. III and IV.

C. The finite bandwidth problem

The goal of this work is to test the agreement between K–K transforms and experimental data using only experimentally determined parameters. This requires some understanding of how artifacts of truncation can affect the K–K predictions. In order to establish some context for distinguishing features inherently related to the causal nature of the data from truncation-associated artifacts, we briefly examine the problem analytically. First, we discuss a model-independent artifact stemming from the Hilbert transform kernel. Next, we examine artifacts associated directly with a particular physical system, the damped resonator. We work with a Hilbert transform pair, the in-phase and quadrature portions of a Lorentzian line, that exhibits resonant behavior and calculate closed-form results for the truncated integrals. These results will illustrate the types of artifacts that can occur with the microbubble data which exhibit analogous resonant structure in its attenuation and dispersion.

1. The Kramers–Kronig transform of a constant over a finite interval

A fundamental model-independent artifact of truncation is due to the kernel of the Kramers–Kronig transform itself. This is seen by performing the integral of the kernel over finite limits (with $\omega_a<\omega<\omega_b$, $\omega_a\geqslant 0$ and $\omega_b<\infty$):

$$I_{\mathrm{KK}}(\omega,\omega_a,\omega_b)\equiv P\int_{\omega_a}^{\omega_b}\frac{d\omega'}{\omega'^2-\omega^2}$$
$$=\frac{1}{2\omega}\left(P\int_{\omega_a}^{\omega_b}\frac{d\omega'}{\omega'-\omega}-\int_{\omega_a}^{\omega_b}\frac{d\omega'}{\omega'+\omega}\right)$$
$$=\frac{1}{2\omega}\ln\left|\frac{\omega_b-\omega}{\omega_b+\omega}\frac{\omega_a+\omega}{\omega_a-\omega}\right|.\qquad(7)$$

When taken over the entire positive axis,

$$I_{\mathrm{KK}}(\omega,0,\infty)=0.\qquad(8)$$

It is the property of the full spectral integration expressed in Eq. (8) that gives one the freedom to add subtraction constants to the numerator [e.g., the $\mathrm{Re}\,\Lambda_n(\omega,\omega_0)$ in Eq. (A4a)]. However, when restricted to a finite interval this integral does not vanish (except at the isolated point $\omega=\sqrt{\omega_a\omega_b}$), and so adding a constant to the numerator does affect the result. Furthermore, the integral diverges as one approaches either end point, $\lim_{\omega\to\omega_a\,\omega_a\neq 0} I_{\mathrm{KK}}(\omega,\omega_a,\omega_b)\to\infty$, and $\lim_{\omega\to\omega_b} I_{\mathrm{KK}}(\omega,\omega_a,\omega_b)\to-\infty$. Figure 1 shows Eq. (7) plotted over an interval from 1 to 15 MHz (the spectral limits of the data to be examined later). These edge-of-bandwidth divergences can be eliminated by replacing an integrand of the form $f(\omega')/(\omega'^2-\omega^2)$ with $[f(\omega')-f(\omega)]/(\omega'^2-\omega^2)$. (This choice also removes the principal value designation from the integral since the integrand is no longer divergent at $\omega=\omega'$.) Still, the use of a subtraction constant can impact the result for integration over a finite interval; thus we may gain finite artifacts but we lose divergent ones. For the Lorentzian pair considered next, we will see that this type of

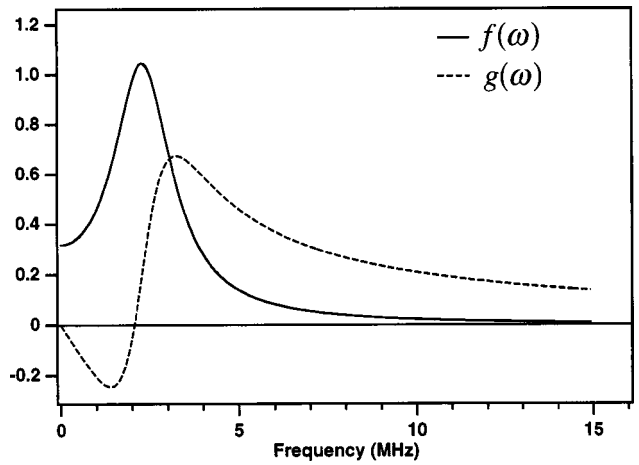


FIG. 2. The Lorentzian Hilbert transform pair from 1 to 15 MHz. The function $f(\omega)$ is given in Eq. (9) and $g(\omega)$ is given in Eq. (10). The half-width, $\Gamma/2\pi$, is 1 MHz and the resonance frequency, $\omega_r/2\pi$, is 2.3 MHz.

subtraction constant {i.e., replacing $f(\omega')/(\omega'^2-\omega^2)$ with $[f(\omega')-f(\omega)]/(\omega'^2-\omega^2)$} is desirable since it both eliminates the infinite artifact and leaves no finite remnant.

2. Artifacts from the application of the truncated relations to a resonant system

Here we present the results of applying the twice-subtracted relations restricted to a finite interval to the Hilbert transform pair defined below. We compare the original functions with their finite-interval twice-subtracted K–K predictions, obtaining closed-form expressions for the differences. These results will aid us in understanding the nature of the truncation artifacts in the K–K transformed microbubble data.

The Hilbert transform pair,

$$f(\omega)=\frac{1}{1+(\omega-\omega_r)^2/\Gamma^2}+\frac{1}{1+(\omega+\omega_r)^2/\Gamma^2} \tag{9}$$

and

$$g(\omega)=\frac{(\omega-\omega_r)/\Gamma}{1+(\omega-\omega_r)^2/\Gamma^2}+\frac{(\omega+\omega_r)/\Gamma}{1+(\omega+\omega_r)^2/\Gamma^2} \tag{10}$$

represent the in-phase and quadrature portions of a Lorentzian line-shape (or the real and imaginary components of the Fourier transform of a damped resonator—see Appendix B). Since $f(\omega)$ and $g(\omega)$ are a Hilbert transform pair (i.e., obey relations with zero subtractions), they also satisfy the dispersion relations with any number of subtractions. We will consider their properties under the twice-subtracted form of the dispersion relation since that is what we will use with the Albunex data considered later. The curves $f(\omega)$ and $g(\omega)$ are plotted in Fig. 2 over the bandwidth of the microbubble data. The values for the half-width ($\Gamma/2\pi=1$ MHz) and resonance frequency ($\omega_r/2\pi=2.3$ MHz) are chosen to give them a rough correspondence to the shape of the Albunex data. We must stress that these two curves are not adequate for making a fit to the Albunex data (this requires more sophisticated functional forms). Rather, the Lorentzian pair was chosen because they represent the components of a resonant structure and yield analytical results in this problem.

The bandlimited estimates of $f(\omega)$ and $g(\omega)$ from the twice-subtracted relations are denoted $f_2^{(\omega_a,\omega_b)}(\omega)$ and $g_2^{(\omega_a,\omega_b)}(\omega)$, respectively, and are given in Eqs. (B15) and (B16). The details of these calculations are given in Appendix B along with an in-depth discussion of the various terms in the results. The differences between our finite bandwidth approximations for $f(\omega)$ and $g(\omega)$ and their true values are given in following:

$$\begin{aligned} f_2^{(\omega_a,\omega_b)}(\omega)-f(\omega) &= -f(\omega)(1-\theta(\omega_a,\omega_b,\omega_r,\Gamma))+g_s(\omega) \\ &\quad\times\varphi(\omega_a,\omega_b,\omega_r,\Gamma)+f(\omega_0)(1-\theta(\omega_a,\omega_b,\omega_r,\Gamma)) \\ &\quad-g_s(\omega_0)\varphi(\omega_a,\omega_b,\omega_r,\Gamma)+(\omega-\omega_0)[f'(\omega_0) \\ &\quad\times(1-\theta(\omega_a,\omega_b,\omega_r,\Gamma))-g_s'(\omega_0)\varphi(\omega_a,\omega_b,\omega_r,\Gamma)], \end{aligned} \tag{11}$$

$$\begin{aligned} g_2^{(\omega_a,\omega_b)}(\omega,\omega_0)-g(\omega) &= -g(\omega)(1-\theta(\omega_a,\omega_b,\omega_r,\Gamma))-f_s(\omega) \\ &\quad\times\varphi(\omega_a,\omega_b,\omega_r,\Gamma)+g(\omega_0)(1-\theta(\omega_a,\omega_b,\omega_r,\Gamma)) \\ &\quad+f_s(\omega_0)\varphi(\omega_a,\omega_b,\omega_r,\Gamma)+(\omega-\omega_0)[g'(\omega_0) \\ &\quad\times(1-\theta(\omega_a,\omega_b,\omega_r,\Gamma))+f_s'(\omega_0)\varphi(\omega_a,\omega_b,\omega_r,\Gamma)]. \end{aligned} \tag{12}$$

[The functions $\theta(\omega_a,\omega_b,\omega_r,\Gamma)$, $\varphi(\omega_a,\omega_b,\omega_r,\Gamma)$, $g_s(\omega)$, and $f_s(\omega)$ are defined by Eqs. (B9)–(B12) in Appendix B.] In Eqs. (11) and (12), we have the closed-form expressions for the artifact that our finite bandwidth restriction causes in our attempt to recover $f(\omega)$ and $g(\omega)$. The last term in each of Eqs. (11) and (12) is the most crucial since each grows linearly away from the subtraction frequency ω_0. The other terms contribute errors that are on the order of or smaller than the values of $f(\omega)$ and $g(\omega)$ themselves.

The nature of these errors is illustrated in Fig. 3. In Fig. 3(a) both the second term and the coefficient of $(\omega-\omega_0)$ in the third term of Eq. (11) are plotted. In Figs. 3(b) and 3(c) the predictions of the truncated Kramers–Kronig relations for $f(\omega)$ are plotted for the choices of ω_0 indicated in the figure. In one case [Fig. 3(b)], the subtraction frequency is chosen so that the third term in Eq. (11) is small and the linearly growing error is suppressed. In the other case [Fig. 3(c)], the subtraction frequency ω_0 is chosen where the coefficient of the $(\omega-\omega_0)$ factor in the third term is large, illustrating the runaway error that can result due to the linear growth away from ω_0. Similar plots for the determination of $g(\omega)$ are shown in Figs. 3(d)–(f). From these examples, it is easily seen that the ability of the truncated relations to recover a reasonable estimate of the desired function is highly dependent on the choice of subtraction frequency ω_0. They also show that reasonably accurate results can be achieved when the large linear artifact is avoided.

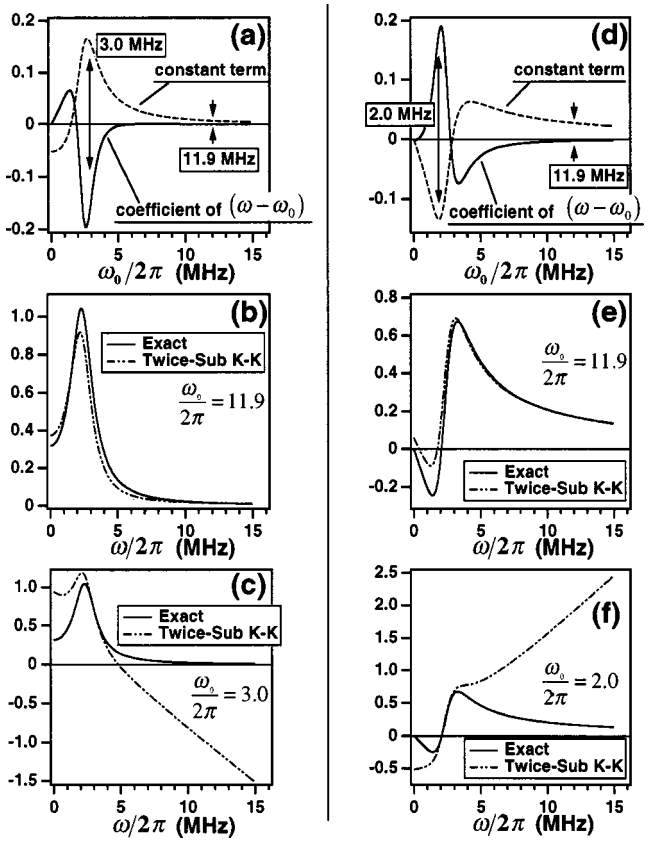


FIG. 3. (a) The two terms contributing to the error in the calculation of $f(\omega)$ using the truncated Kramers–Kronig relation with two subtractions. The prediction for $f(\omega)$ is calculated using the subtraction frequencies indicated in this panel. (b) The exact $f(\omega)$ along with the truncated Kramers–Kronig prediction using a subtraction frequency of 11.9 MHz where the error terms are small. (c) The prediction using the subtraction frequency 3.0 MHz where there is a large deviation from the desired result. The large value of the linear term's coefficient causes the runaway artifact. (d) The two terms contributing to the error in the calculation of $g(\omega)$ using the truncated Kramers–Kronig relation with two subtractions. The prediction for $g(\omega)$ is calculated using the subtraction frequencies indicated in this panel. (e) The exact $g(\omega)$ along with the truncated Kramers–Kronig prediction using a subtraction frequency of 11.9 MHz where the error terms are small. (f) The prediction using the subtraction frequency 2.0 MHz where there is a large deviation from the desired result.

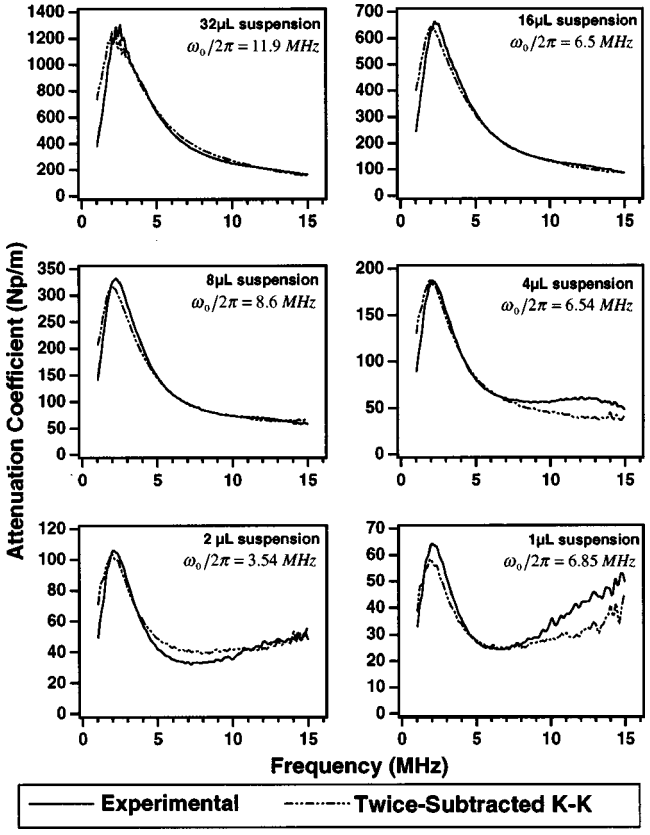


FIG. 4. Attenuation coefficient results from the twice-subtracted K–K relation [Eq. (13)] for the six Albunex suspensions. Both the K–K predictions and experimental data are shown. The subtraction frequencies used in the calculations are indicated in each panel of the figure. The suspensions are labeled by the amount of Albunex they contain (1–32 μL).

III. METHODS

A. Description of microbubble suspensions and data

Albunex, a first-generation contrast agent for echocardiography, contains air microbubbles that are stabilized by encapsulating shells of the denatured protein albumin. The majority of the microbubble volume comes from shells with radii from 2 to 5 μm. These suspensions are of interest in this work because sufficiently diluted Albunex behaves like a notch filter for ultrasound with an attenuation peak just above 2 MHz. The resonant mode corresponds to a uniform radial (monopolar) oscillation of the encapsulated bubble. The experimental data comes from measurements performed in our laboratories (Washington U. and Mallinckrodt) using the broadband ultrasonic spectroscopy technique.[40,41] The phase velocity data used in this work were reported in an earlier paper.[40] The attenuation coefficient data used here were derived from the same waveforms used to calculate phase velocity. These specific attenuation data have not been published but are consistent with our previous work on the attenuation properties of Albunex that reported data from earlier experiments.[42] The six suspensions we examined consisted of various amounts of Albunex diluted into 8 mL of Isoton (buffered saline). We will refer to the suspensions by the amounts of Albunex they contain, which run from 1 μL up to 32 μL in powers of 2. Each data set consists of 229 discrete points uniformly covering the spectrum from 1 to 15 MHz. The frequency spacing between successive data points is $\Delta\omega/2\pi=0.061\,035$ MHz.

B. Numerical methods

The integrals were approximated as Riemann sums by calculating the integrand for each point in a given data set, adding these values together, and multiplying the resulting sum by the frequency interval $\Delta\omega$. At points where the integrands [see Eqs. (13)–(16)] were indeterminate (i.e., of the form 0/0 when $\omega'=\omega$ or ω_0) l'Hopital's rule was used to obtain its proper limiting form. The derivative terms were approximated using finite differences of the form $\alpha'(\omega)\rightarrow\alpha'[n]=(\alpha[n+1]-\alpha[n-1])/2\Delta\omega$, except at the endpoints where single-sided differences were used.

IV. RESULTS

In this section we present the results of the interval-restricted Kramers–Kronig calculations using attenuation co-

efficient and phase velocity data from Albunex suspensions. The expansion form of the twice-subtracted K–K predictions [Eqs. (3) and (5)] are used as starting points. Lower-order [once -subtracted K–K in the case of $\alpha(\omega)$] calculations are also performed.

A. Attenuation coefficient results

The twice-subtracted K–K predictions for the attenuation coefficient were calculated using the finite-interval form of Eq. (3),

$$\alpha(\omega)=\left(\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\omega'/c(\omega')-\omega/c(\omega)}{\omega'-\omega}d\omega'+\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\omega'/c(\omega')+\omega/c(\omega)}{\omega'+\omega}d\omega'\right)$$
$$+\left(\alpha(\omega_0)-\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\omega'/c(\omega')-\omega_0/c(\omega_0)}{\omega'-\omega_0}d\omega'-\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\omega'/c(\omega')+\omega_0/c(\omega_0)}{\omega'+\omega_0}d\omega'\right)$$
$$+(\omega-\omega_0)\left(\frac{d}{d\omega}\alpha(\omega)\big|_{\omega=\omega_0}-\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\mathrm{Im}\,\Lambda_2(\omega',\omega_0)\,d\omega'-\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\mathrm{Im}\,\Lambda_2(-\omega',\omega_0)\,d\omega'\right). \quad (13)$$

The comparisons between the experimental curves and predictions using Eq. (13) are shown in Fig. 4. The K–K curves are functions of the subtraction frequency ω_0, and for these calculations the ω_0's were chosen by looking for the best agreement between K–K and the data over the entire experimental spectrum. The ω_0's used are indicated in each panel of Fig. 4. (The role of the ω_0's in these calculations is explored further in Sec. V.)

Figure 5 compares the experimental data with the finite interval prediction of the once-subtracted relation,

$$\alpha(\omega)=\left(\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\omega'/c(\omega')-\omega/c(\omega)}{\omega'-\omega}d\omega'+\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\omega'/c(\omega')+\omega/c(\omega)}{\omega'+\omega}d\omega'\right)$$
$$+\left(\alpha(\omega_0)-\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\omega'/c(\omega')-\omega_0/c(\omega_0)}{\omega'-\omega_0}d\omega'-\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\omega'/c(\omega')+\omega_0/c(\omega_0)}{\omega'+\omega_0}d\omega'\right), \quad (14)$$

which differs from Eq. (13) by the omission of the linear factor. Unlike the case with the twice-subtracted results, the shape and magnitude (minimum-to-maximum variation) of these once-subtracted curves are independent of the subtraction frequency.

B. Phase velocity results

For the phase velocity, the finite interval form of the twice-subtracted relation [Eq. (5)] is given by

$$\omega\left(\frac{1}{c(\omega)}-\frac{1}{c(\omega_0)}\right)=\left(\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\alpha(\omega')-\alpha(\omega)}{\omega'-\omega}d\omega'-\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\alpha(\omega')-\alpha(\omega)}{\omega'+\omega}d\omega'\right)$$
$$-\left(\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\alpha(\omega')-\alpha(\omega_0)}{\omega'-\omega_0}d\omega'-\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\alpha(\omega')-\alpha(\omega_0)}{\omega'+\omega_0}d\omega'\right)$$
$$+(\omega-\omega_0)\left(\omega_0\frac{d}{d\omega}\frac{1}{c(\omega)}\bigg|_{\omega=\omega_0}+\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\mathrm{Re}\,\Lambda_2(\omega',\omega_0)\,d\omega'+\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\mathrm{Re}\,\Lambda_2(-\omega',\omega_0)\,d\omega'\right). \quad (15)$$

Figure 6 compares the dispersion curves [in the form $1/c(\omega)-1/c(\omega_0)$] calculated from the twice-subtracted relation [Eq. (15) divided by ω] and the experimental data, using the subtraction frequencies (ω_0's) indicated in the various panels. The ω_0's were chosen to obtain the best match between the predictions and experimental data for each respective suspension. (The role of the ω_0's in these calculations is explored further in Sec. V.)

If the linear term on the right-hand side of Eq. (15) is ignored, the remaining four terms in the first two brackets have the form of the once-subtracted relation for the phase velocity. However, we found that these four integrals produced estimates of $\omega(1/c(\omega)-1/c(\omega_0))$ and not of $\omega/c(\omega)-\omega_0/c(\omega_0)$ as one would expect if the once-subtracted relation held. So we refer to the expression

$$\omega\left(\frac{1}{c(\omega)}-\frac{1}{c(\omega_0)}\right)$$
$$=\left(\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\alpha(\omega')-\alpha(\omega)}{\omega'-\omega}d\omega'-\frac{1}{\pi}\times\int_{\omega_a}^{\omega_b}\frac{\alpha(\omega')-\alpha(\omega)}{\omega'+\omega}d\omega'\right)-\left(\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\alpha(\omega')-\alpha(\omega_0)}{\omega'-\omega_0}\times d\omega'-\frac{1}{\pi}\int_{\omega_a}^{\omega_b}\frac{\alpha(\omega')-\alpha(\omega_0)}{\omega'+\omega_0}d\omega'\right) \quad (16)$$

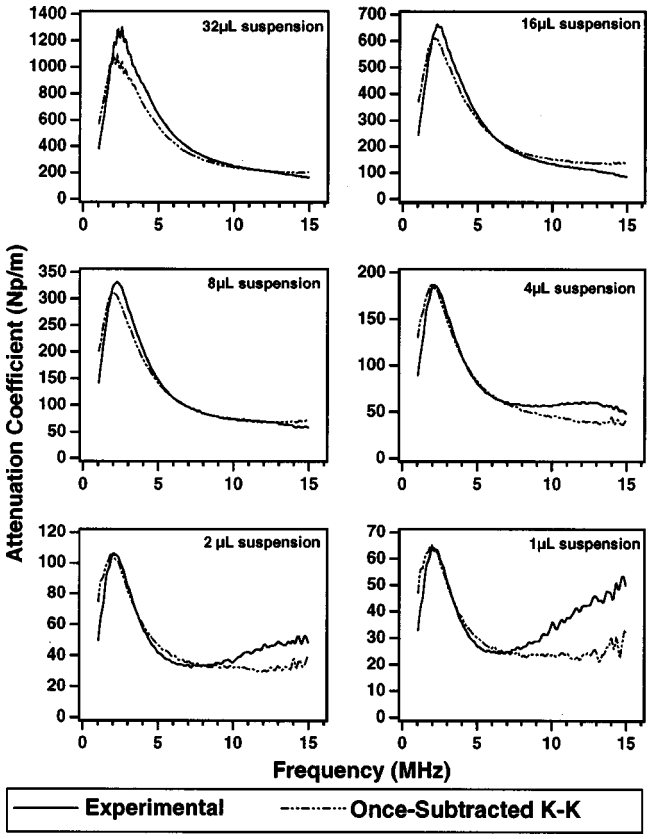


FIG. 5. Attenuation coefficient results from the once-subtracted K–K relation [Eq. (14)] for the six Albunex suspensions. Both the K–K predictions and experimental data are shown. The shapes of the predictions are independent of the subtraction frequency.

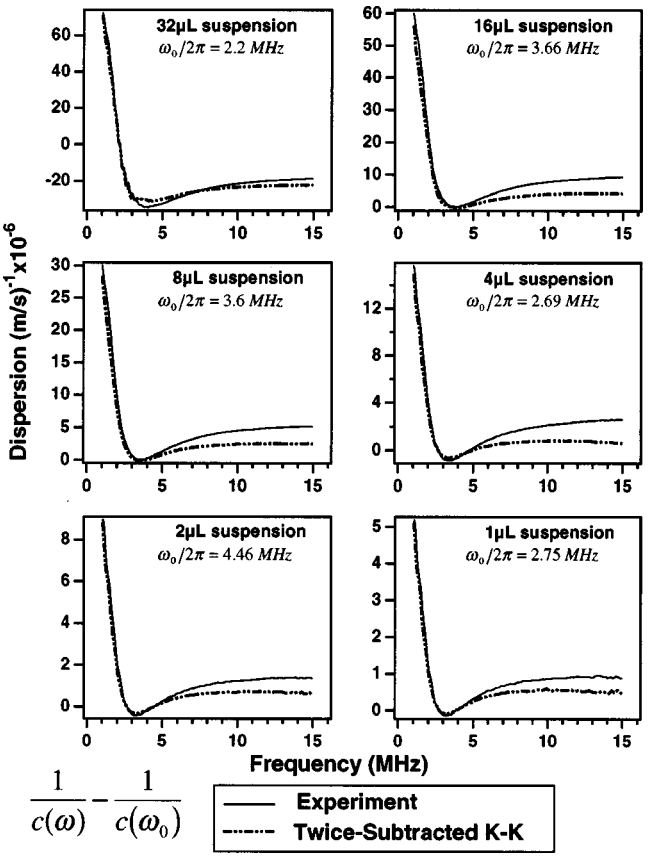


FIG. 6. The dispersion results [in the form $1/c(\omega)-1/c(\omega_0)$] from the twice-subtracted K–K relation [Eq. (15) divided by ω] for the six Albunex suspensions. Both the K–K predictions and experimental data are shown. The subtraction frequencies used in the calculations are indicated in the figure. The shapes of the experimental curves are independent of the subtraction frequency. (Note that the zero points of dispersion are set by the choice of ω_0 for the K–K calculation and are not inherent features of the experimental data.)

[Eq. (15) with the linear term removed] as the ''lower-order'' relation as opposed to ''once-subtracted'' which would be incorrect in this case. Figure 7 is the comparison of the lower-order predictions [Eq. (16) divided by ω] of $1/c(\omega)-1/c(\omega_0)$ and experimental curves. As before, the ω_0's were chosen to obtain the best match between the predictions and the experimental data.

V. DISCUSSION

A. Morphology of the predicted curves

The curves produced with the twice-subtracted attenuation predictions and both orders of dispersion predictions have shapes that are dependent on the choice of ω_0. The best choices for ω_0 were judged to be those that achieved the closest match in shape between the predicted and experimental curves. As can be seen in Figs. 4, 6, and 7, these K–K calculations reproduced the desired shapes quite well. The once-subtracted $\alpha(\omega)$ predictions stand alone for they are the only K–K results whose predicted curves have ω_0-independent shapes. The variations of these curves are determined solely from the Hilbert transform of $\omega/c(\omega)$. Since these predictions are not colored by the presence of the ''tuning parameter'' ω_0, they provide a stringent test on the ability of the Hilbert transforms to generate the morphology exhibited by the target data. For the 1-, 2-, 4-, and 8-μL suspensions the once- and twice-subtracted predictions perform to roughly the same level of accuracy for $\alpha(\omega)$, as can be seen by comparing Figs. 4 and 5. However, for the two most concentrated suspensions, 16 and 32 μL, the once-subtracted curves do not track the attenuation data as well at the high-frequency end. For the twice-subtracted predictions, the linear term modifies the Hilbert transform curves delivering the accurate matches at the high-frequency end for these same suspensions. The $\alpha(\omega)$ determinations for the 16- and 32-μL data sets are the only clear cases where the linear term markedly improved the accuracy of the predictions over the lower-order results.

The predictions and experimental data can be compared more quantitatively by looking at the magnitudes of the variations in the assorted curves. For the attenuation coefficient we use the parameter $\Delta\alpha=\max[\alpha(\omega)]-\min[\alpha(\omega)]$. The three values of $\Delta\alpha$ (two K–K and one experimental) for each suspension are shown in Fig. 8(a). In Fig. 8(b) the ratios of the predictions to the experimental values are compared. As seen in Fig. 8(b), accuracy of the two types of K–K predictions for the four most dilute suspensions are roughly similar. For the two most concentrated suspensions, the decreasing accuracy of the once-subtracted prediction is evident.

For the phase velocity, both the lower-order and twice-subtracted calculations have minimum-to-maximum varia-

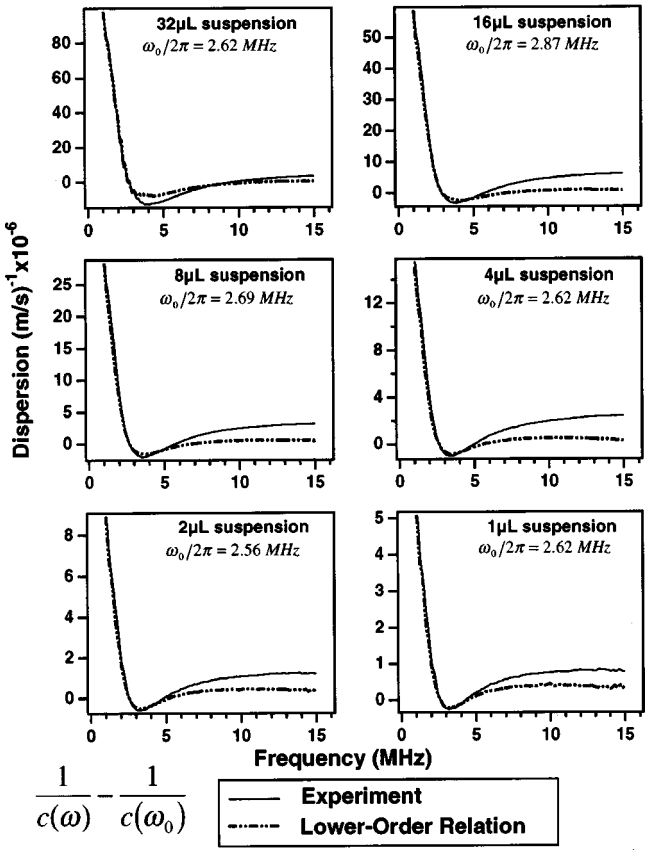


FIG. 7. The dispersion results (in the form $1/c(\omega)-1/c(\omega_0)$) from the lower-order K–K relation [Eq. (16) divided by ω] for the six Albunex suspensions. Both the K–K predictions and experimental data are shown. The subtraction frequencies used in the calculations are indicated in the figure. The shapes of the experimental curves are independent of the subtraction frequency. (Note that the zero points of dispersion are set by the choice of ω_0 for the K–K calculation and are not inherent features of the experimental data.)

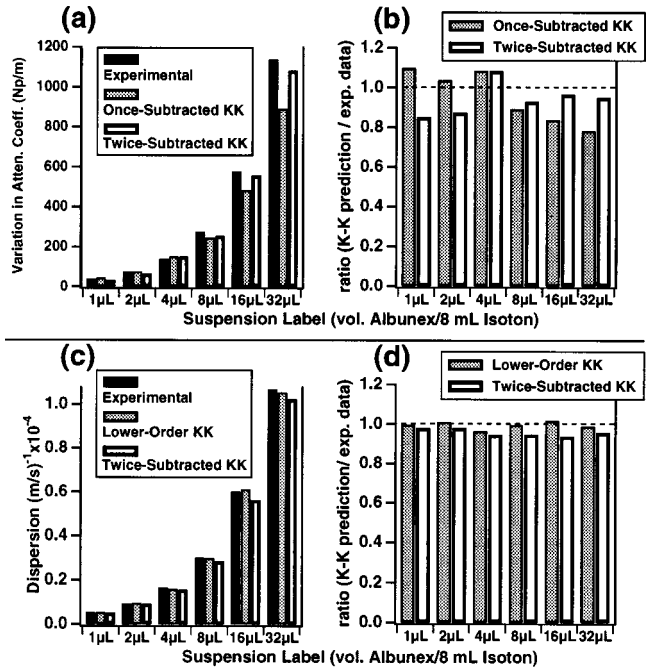


FIG. 8. A comparison of the Kramers–Kronig predictions and the experimental values for the variation in the attenuation coefficient and dispersion curves. (a) The comparison of $\Delta\alpha$ for the data, the once-subtracted, and twice-subtracted predictions across the six Albunex suspensions. These quantities are calculated from the curves shown in Figs. 4 and 5. (b) The ratios of the respective predictions for $\Delta\alpha$ with the experimental values. (c) The comparisons of the magnitude of dispersion [defined as $\max(1/c(\omega)-1/c(\omega_0))-\min(1/c(\omega)-1/c(\omega_0))$] for the data, the lower-order, and the twice-subtracted predictions across the six Albunex suspensions. (The magnitude of dispersion of the experimental data is independent of ω_0, unlike the K–K predicted values). These quantities are calculated from the curves shown in Figs. 6 and 7. (d) The ratio of the respective predictions with the experimental value for the magnitudes of dispersion.

tions (i.e., magnitudes of dispersion) that closely match the experimental data. These magnitudes of dispersion are plotted in Fig. 8(c) and the predicted-to-experimental ratios are given in Fig. 8(d). Both K–K orders perform well across the six suspensions, with the lower-order values exhibiting slightly better agreement than the twice-subtracted ones. In terms of absolute phase velocity, the differences between the data and both orders of predictions are $\leqslant 1\%$ at each frequency for all six suspensions.

B. Role of subtraction frequency ω_0 in twice-subtracted predictions

To achieve reasonable results for both $\alpha(\omega)$ and $1/c(\omega)-1/c(\omega_0)$ with the twice-subtracted calculations, the choice of ω_0 is critical. For each suspension, there were only a few choices of subtraction frequency that resulted in close agreement with the experimental data. For most other choices of ω_0, the twice-subtracted predictions exhibited some manifestation of the linear artifact of the type described in the study of the Lorentz model. When reasonable agreement was achieved, the coefficient of the linear terms in Eqs. (13) and (15), respectively, always took on values among the smallest available. This is illustrated in Fig. 9 using the data for the 8-μL Albunex suspension, where the consequences of poor choices for ω_0 in the twice-subtracted calculations are shown. In Fig. 9(a), the coefficient of the linear term,

$$\frac{d}{d\omega}\alpha(\omega)\big|_{\omega=\omega_0} - \frac{1}{\pi}\int_{\omega_a}^{\omega_b} \mathrm{Im}\,\Lambda_2(\omega',\omega_0)\,d\omega' - \frac{1}{\pi}\int_{\omega_a}^{\omega_b} \mathrm{Im}\,\Lambda_2(-\omega',\omega_0)\,d\omega', \tag{17}$$

from the twice-subtracted relation for $\alpha(\omega)$ [Eq. (13)] is plotted. Three different points on the curve are indicated. Two of the choices, 2.0 and 4.2 MHz, correspond to points near extrema of this curve. Figures 9(b) and (c) display the results of using either of these choices as $\omega_0/2\pi$. The runaway linear artifact is clear. Using the point at 8.6 MHz, near the indicated zero crossing of Eq. (17), resulted in an accurate prediction. This is the subtraction frequency used to calculate the 8-μL curve in Fig. 4. When Eq. (17) is sufficiently far away from zero, the linear artifact dominates the results.

The right-hand side of Fig. 9 illustrates the linear artifact for the dispersion calculations. In Fig. 9(d), the coefficient of the linear term,

$$\omega_0 \frac{d}{d\omega}\frac{1}{c(\omega)}\bigg|_{\omega=\omega_0} + \frac{1}{\pi}\int_{\omega_a}^{\omega_b} \mathrm{Re}\,\Lambda_2(\omega',\omega_0)\,d\omega' + \frac{1}{\pi}\int_{\omega_a}^{\omega_b} \mathrm{Re}\,\Lambda_2(-\omega',\omega_0)\,d\omega', \tag{18}$$

from Eq. (15) is plotted. Two extreme values for the coefficient are pointed out and the consequences of using those subtraction frequencies (2.0 and 5.25 MHz) for the twice-subtracted calculations are shown in Figs. 9(e) and (f). For the best agreement (as shown as the 8-μL curve in Fig. 6) the 3.56-MHz point, near a zero-crossing of the linear coefficient, in the top panel was used. As before, when the coefficient [Eq. (18)] is sufficiently small, the linear artifact is avoided.

For each of the six suspensions, the best prediction of the $c(\omega)$-to-$\alpha(\omega)$ relation involved a different subtraction frequency than the best $\alpha(\omega)$-to-$c(\omega)$ calculation. This is not surprising or troubling as it is also a feature of the Lorentzian model. As just discussed, the criterion for choosing ω_0 in the twice-subtracted relations is to minimize/suppress the coefficient of the linear term. Since the linear terms in the $c(\omega)$-to-$\alpha(\omega)$ and $\alpha(\omega)$-to-$c(\omega)$ relations have different zero-crossings, achieving the best agreements for any single suspension requires using different values for ω_0 in the two respective transformations. For the 8-μL suspension, this can be seen by comparing the two curves plotted in Figs. 9(a) and (d). The low-frequency zero-crossing in Fig. 9(a) occurs near 2.5 MHz while in Fig. 9(d) the first zero-crossing is at 3.6 MHz. In the Lorentzian model, this same feature can be seen by comparing the linear coefficient terms from Figs. 3(a) and (d). In this case the low-frequency zero-crossings in one of the curves correspond to extrema in the other curve. Instead of optimizing ω_0 individually for each transformation, one could also try to choose a single ω_0 for both $c(\omega)$-to-$\alpha(\omega)$ and $\alpha(\omega)$-to-$c(\omega)$ transformations of a given suspension's data. In this case, accurate (but not optimum) results are also possible. With the 8-μL data, there are regions where both coefficients are relatively small, particularly near the high-frequency end of the bandwidth. This is also true of the other five suspensions. So, choosing a single ω_0 for each suspension where both of the linear coefficients are small should yield reasonably accurate results for both of the twice-subtracted transformations.

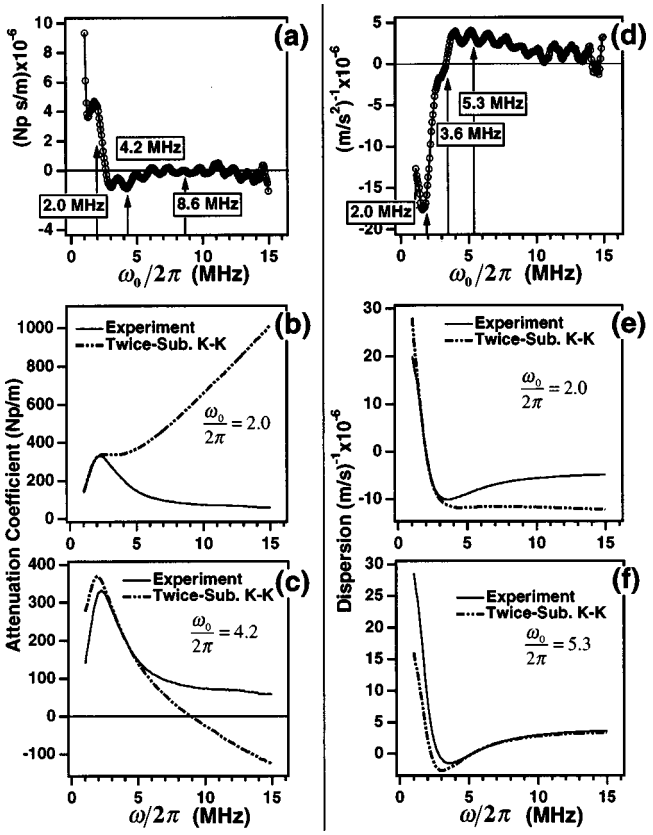


FIG. 9. Illustrations of the runaway linear artifact in the twice-subtracted K–K calculations. (a) The coefficient of the linear term in the twice-subtracted relation for the attenuation coefficient [Eq. (17)] for the 8-μL suspension. The values at three subtraction frequencies, 2.0, 4.2, and 8.6 MHz, are indicated. At 8.6 MHz, the coefficient was near zero and this value was used for calculating the result shown in Fig. 4. (b) The result of using 2.0 MHz (where the coefficient in the top panel has a relatively large positive value) as the subtraction frequency for the attenuation coefficient calculation. (c) The result of using 4.2 MHz (where the coefficient in the top panel is near its most negative value) as the subtraction frequency. (d) The coefficient of the linear term in the twice-subtracted phase velocity relation [Eq. (18)] for the 8-μL suspension. The values at three subtraction frequencies, 2.0, 3.6, and 5.25 MHz, are indicated. At 3.6 MHz, the coefficient was near zero and this value was used for calculating the result shown in Fig. 7. (e) The result of using 2.0 MHz (where the coefficient in the top panel is near its most negative value) as the subtraction frequency for the phase velocity calculation. (f) The result of a calculation using 5.25 MHz (where the coefficient in the top panel has a relatively large positive value) as the subtraction frequency.

C. General remarks

The evidence is suggestive but not conclusive on the necessity of the linear terms in the truncated relations with these data. The troubling aspect of the linear terms in the twice-subtracted relations are the strong artifacts they can introduce. In all cases the consistent feature of choosing the subtraction frequency was that it lie near a zero of the coefficient of the linear term. It is clear that the linear term must be suppressed by keeping the coefficient small. But it is not known that the coefficient is any more valid when it is small. So it remains to be seen if the improvement in the two attenuation curves (16 and 32 μL) was a causally linked result or an artifactual accident. It seems to be a proper improvement but conclusive proof is still lacking. Unfortunately, the Lorentzian model analysis could provide no guidance on the proper size of the linear term. Because the Lorentzian pair satisfy the zeroth-order relations, the linear terms in their transformations should be identically zero. Thus any nonzero value for the linear coefficient for the model is an artifact. Further analysis of these data, perhaps in a future study, could certainly provide more insight on this issue.

This work supports the validity of the twice-subtracted form of the phase velocity relation for the Albunex data. This was clearly indicated by the calculations themselves, since the bandlimited form of the terms

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\alpha(\omega')-\alpha(\omega)}{\omega'-\omega}\,d\omega' - \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\alpha(\omega')-\alpha(\omega_0)}{\omega'-\omega_0}\,d\omega'$$

returned estimates of $\omega(1/c(\omega)-1/c(\omega_0))$. If the proper once-subtracted relations were valid, then the quantity $\omega/c(\omega)-\omega_0/c(\omega_0)$ should result. In the form of the twice-subtracted phase velocity relation used in this work, the difference between $\omega(1/c(\omega)-1/c(\omega_0))$ and $\omega/c(\omega)-\omega_0/c(\omega_0)$ has been absorbed into the linear coefficient. From our experience with these Albunex data sets, the proper once-subtracted relation for $c(\omega)$ does not hold over the finite bandwidth.

Our results also provide evidence in support of the ''log-of-the-system-response/method-of-subtractions'' route to the K–K relations. The strength of this method lies with its generality—one does not need to know a constitutive relation or some other specific attribute of the system under investigation, only that the general requirements of causality, passivity, and linearity are met. The complex wave number parametrization is a general model that is natural for use in wave propagation problems in linear media. Measurement techniques such as the broadband ultrasonic spectroscopic methods are built on the foundations of the same general model, where the magnitude and phase of the transfer function are parametrized using the attenuation coefficient and phase velocity. Using a specific constitutive relation in a K–K derivation may be desirable in certain situations. However, when relationships between $\alpha(\omega)$ and $c(\omega)$ are desired, the constitutive approach usually produces more complicated relations than the general transfer function method.

In actual practice, one may want to determine the attenuation coefficient from dispersion data alone, and vice versa. Validating specific algorithms for performing these tasks is beyond the scope of this work, but using the lower-order relations may provide a reasonable estimate of the shape of the desired curve without knowing any derivatives. All of the dispersion and most of the attenuation data for the Albunex suspensions were well predicted by the lower-order relations. One could then devise a criterion for choosing the subtraction frequency based on the result of the calculations (e.g., a magnitude of dispersion halfway between the extreme values, or some physical constraint on the system). One might suggest using the outcome of a lower-order calculation to provide the necessary derivative term for the twice-subtracted relation. However, it is simple to show that the coefficient of the linear term is identically zero in this case, reducing the twice-subtracted relation down to the lower-order.

Although we focused on truncation as the source of the errors and artifacts, one could also view them as arising from the missing information beyond our measurement spectrum. The outcome of the Lorentzian model will remain the same, but perhaps some further physical insight could be gained from this viewpoint. With the agreement we found for the Albunex data, it seems that most of the information necessary for getting accurate results was within the measurement spectrum. Thus we chose to view the problem entirely from within the known portion of the spectrum.

The work on the Lorentzian model was key in establishing confidence in these results. It showed that the artifacts of truncation would not necessarily overwhelm or dominate the results, and demonstrated the inherent problem of the linear artifact in the truncated twice-subtracted relations. It also pointed us toward the ''Hilbert transform+power series expansion'' form of the relations which has proven to be a very useful way of viewing the K–K problem for a resonant system over a finite interval.

VI. CONCLUSION

In this study, we applied the fully expressed Kramers–Kronig relations over a finite bandwidth to experimental data exhibiting resonant behavior. We found both the twice-subtracted and lower-order K–K transforms produced predictions that were highly consistent with the Albunex data sets. To varying degrees the transforms reproduced both the size and shape of the target phase velocity and attenuation coefficient data across all six Albunex suspensions. The twice-subtracted K–K relations produced accurate estimates for both the attenuation coefficient and dispersion for all 12 data sets. Lower-order Kramers–Kronig relations also produced good results over the finite spectrum for most of the data. In 2 of the 12 cases, the twice-subtracted relations tracked the data markedly better than the lower-order predictions. This work provided a strict test of these relations as all parameters used in the calculations were from within the data sets. These results also offer evidence in favor of the validity of the subtracted forms of the Kramers–Kronig relations derived from the logarithm of the transfer function. In order to distinguish between artifactual and causal influences on the K–K outcomes, we also explored the finite-interval problem theoretically, using an analytical model to gain an understanding of the consequences of truncation. These theoretical results demonstrate that good agreement can be achieved and also pointed out the potentially dominant linear artifact inherent in the twice-subtracted K–K relation. The model also illustrates the suppression of the artifact by the proper choice of subtraction frequency, a method that also served to minimize artifactual influences in the microbubble calculations.

APPENDIX A: DERIVATION OF TWICE-SUBTRACTED KRAMERS–KRONIG RELATIONS

The derivation of Kramers–Kronig relations for a susceptibility or transfer function of a passive linear medium is straightforward. For an extended system the following general relation expresses the propagation of a disturbance:

$$f(x_2,t)=\int_{-\infty}^{\infty} h(x_2-x_1,t-t')g(x_1,t')\,dt', \tag{A1}$$

where $h(D,\tau)$ is the impulse response for a material of thickness D, $g(x_1,t)$ is the stimulus, and $f(x_2,t)$ is the response. Using Titchmarsh's theorem,[4,26] one can show that the impulse response $h(D,\tau)$ has a Fourier transform $H(D,\omega)=A(D,\omega)+iB(D,\omega)$ whose real and imaginary parts are Hilbert transforms of one another,

$$A(D,\omega)=\frac{1}{\pi}P\int_{-\infty}^{\infty}\frac{B(D,\omega')}{\omega'-\omega}d\omega'$$

and

$$B(D,\omega) = -\frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{A(D,\omega')}{\omega' - \omega} d\omega'.$$

This conclusion stems from causality, $h(D,\tau<0)=0$, and the finite energy condition, $\int_{-\infty}^{\infty}|h(D,t)|^2 dt<\infty$.

1. Analytic properties of the complex wave number and the method of subtractions

If only the primary disturbance is considered (that is, reverberations are ignored) the transfer function can be written in its magnitude and phase form in terms of the attenuation coefficient $\alpha(\omega)$ and phase velocity $c(\omega)$,

$$H(D,\omega) = \exp(iK(\omega)D) \tag{A2a}$$

$$= \exp\left[\left(-\alpha(\omega) + i\frac{\omega}{c(\omega)}\right)D\right] \tag{A2b}$$

and

$$\frac{\ln H(D,\omega)}{D} = iK(\omega) = -\alpha(\omega) + i\frac{\omega}{c(\omega)}. \tag{A2c}$$

Here it is clear that the complex wave number $K(\omega)$ is a property of the bulk material of the sample and should be independent of the sample dimensions (a consequence of the linear assumption). From this point on, we chose to work with $iK(\omega)$ instead of $K(\omega)$ to preserve the conventional parity arrangements of the real and imaginary parts in the dispersion relations. We define the symbol $\gamma(\omega) \equiv iK(\omega)$ for use in the remainder of Appendix A. Note that there are any number of functions that could be used for the phase term and here the choice of $\arg H(D,\omega) \equiv [\omega/c(\omega)]D$ is made which avoids the problem of branch cuts. The task is to relate the attenuation coefficient and the phase velocity by the Hilbert transforms. This will be accomplished by forming a new function $\Lambda_n(\omega,\omega_0)$ in terms of $\gamma(\omega)$ that obeys condition (*ii*) of Titchmarsh's theorem:[27] (*ii*) if a square-integrable function $S(z=\omega+iy)$ is analytic in the complex plane for $y>0$, approaches $S(\omega)$ almost everywhere as $y\to 0$, and is square integrable along every line in the upper-half plane parallel to the real axis, $\int_{-\infty}^{\infty}|S(\omega+iy)|^2 d\omega<\infty$, then its real and imaginary parts obey the Hilbert transforms on the ω axis. Since $H(D,\omega)$ is analytic everywhere in the upper-half plane, $\gamma(z=\omega+iy)=[\ln H(D,z)]/D$ is as well except where $H(D,z)$ is zero. Nussenveig[27] and others[5] have concluded that there are no zeroes of $H(D,z)$ in the upper-half plane for finite z (minimum phase condition) due to its specific dependence on the arbitrary thickness variable D. The remaining obstacle is the behavior of the transfer function as $|z|\to\infty$. Since the transfer function is square integrable, $\int_{-\infty}^{\infty}|H(D,\omega+iy)|^2 d\omega<M$, it goes to zero when $|z|\to\infty$ on the real axis and in the upper-half plane, that is, $\lim_{\omega\to\pm\infty, y\geq 0} H(D,\omega+iy)\to 0$. If $H(x,z)\sim 1/z^n$ or $H(x,z)\sim\exp(-|z|^n)$ as $|z|\to\infty$, then $\gamma(z)\sim n\ln z$ or $\gamma(z)\sim z^n$. Thus $\gamma(z)$ is divergent and not square integrable. This difficulty is overcome with the method of subtractions.[27] If $\gamma(\omega)$ diverges as ω^n (or slower), a square integrable function $\Lambda_n(\omega,\omega_0)$ can be formed from $\gamma(\omega)$:

$$\Lambda_n(\omega,\omega_0) = \frac{\gamma(\omega) - \gamma(\omega_0) - (d/d\omega)\gamma(\omega)\big|_{\omega=\omega_0}(\omega-\omega_0) - \cdots - (d^{n-1}/d\omega^{n-1})\gamma(\omega)\big|_{\omega=\omega_0}(\omega-\omega_0)^{n-1}/(n-1)!}{(\omega-\omega_0)^n}. \tag{A3}$$

This new function $\Lambda_n(z,\omega_0)$ is analytic everywhere $\gamma(z)$ is. It is also convergent to zero as $|z|\to\infty$ in the upper-half plane and square integrable. Therefore $\Lambda_n(\omega,\omega_0)$ satisfies condition (*ii*) of Titchmarsh's theorem and consequently the real and imaginary components of $\Lambda_n(\omega,\omega_0)$ form a Hilbert transform pair. The resulting expressions are known as dispersion relations with n subtractions and the point ω_0 is called the substraction frequency. Using $P\int_{-\infty}^{\infty} d\omega'/(\omega'-\omega)=0$, a constant can be added to each respective numerator to keep the integrands finite at $\omega=\omega'$ and remove principal value designation. The relations are given as follows:

$$\mathrm{Im}\,\Lambda_n(\omega,\omega_0) = -\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\mathrm{Re}\,\Lambda_n(\omega',\omega_0)-\mathrm{Re}\,\Lambda_n(\omega,\omega_0)}{\omega'-\omega}d\omega' \tag{A4a}$$

$$= -\frac{1}{\pi}\int_{0}^{\infty}\frac{\mathrm{Re}\,\Lambda_n(\omega',\omega_0)-\mathrm{Re}\,\Lambda_n(\omega,\omega_0)}{\omega'-\omega}d\omega' + \frac{1}{\pi}\int_{0}^{\infty}\frac{\mathrm{Re}\,\Lambda_n(-\omega',\omega_0)-\mathrm{Re}\,\Lambda_n(\omega,\omega_0)}{\omega'+\omega}d\omega' \tag{A4b}$$

$$\mathrm{Re}\,\Lambda_n(\omega,\omega_0) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\mathrm{Im}\,\Lambda_n(\omega',\omega_0)-\mathrm{Im}\,\Lambda_n(\omega,\omega_0)}{\omega'-\omega}d\omega' \tag{A5a}$$

$$= \frac{1}{\pi}\int_{0}^{\infty}\frac{\mathrm{Im}\,\Lambda_n(\omega',\omega_0)-\mathrm{Im}\,\Lambda_n(\omega,\omega_0)}{\omega'-\omega}d\omega' - \frac{1}{\pi}\int_{0}^{\infty}\frac{\mathrm{Im}\,\Lambda_n(-\omega',\omega_0)+\mathrm{Im}\,\Lambda_n(\omega,\omega_0)}{\omega'+\omega}d\omega'. \tag{A5b}$$

The second integrals in Eqs. (A4b) and (A5b) are the negative frequency portion of the original integrals transformed to positive frequency. Once a minimum number of subtractions is found to be sufficient, dispersion relations of higher order are also valid. For the acoustics case, it is not clear on physical grounds what the asymptotic forms of $[\ln H(D,\omega)]/D = -\alpha(\omega) + i\omega/c(\omega)$ should be. This question, although important from a strictly mathematical point of view, may be irrelevant since the classical concepts of continuum mechanics will become invalid in the high-frequency

limit. Ignoring this for the moment, the Paley–Wiener criterion[43] puts a constraint (arising from causality) on the divergence of the attenuation coefficient, restricting it to a slower than linear rise with frequency in the high-ω limit. Weaver and Pao[5] reach a similar conclusion. If the transfer function is treated as a distribution rather than a point function, a ''generalized'' Paley–Wiener criterion holds that is less restrictive.[44] In any case $\alpha(\omega)$ must diverge in the high-frequency limit for $H(D,\omega)$ to be square integrable. One subtraction is then not sufficient to satisfy the square-integrability condition. Based on this reasoning, the dispersion relations with two subtractions should be the lowest-order transforms valid for relating $\alpha(\omega)$ and $c(\omega)$, or at least for calculating $c(\omega)$ from $\alpha(\omega)$. However, when restricting the integrals to a finite bandwidth, a lower-order relation might hold to a reasonable degree since convergence problems are avoided.

2. Kramers–Kronig relations for the phase velocity and attenuation coefficient with two subtractions

Using Eqs. (A4b) and (A5b) and the parity property $\gamma(-\omega)=\gamma^*(\omega)$, the twice-subtracted relations for the attenuation coefficient and phase velocity are obtained:

$$\alpha(\omega)=\alpha(\omega_0)+(\omega-\omega_0)\alpha'(\omega_0)-\frac{(\omega-\omega_0)^2}{\pi}\int_0^\infty\frac{\mathrm{Im}\,\Lambda_2(\omega',\omega_0)-\mathrm{Im}\,\Lambda_2(\omega,\omega_0)}{\omega'-\omega}d\omega'$$
$$-\frac{(\omega-\omega_0)^2}{\pi}\int_0^\infty\left(\frac{\omega'/c(\omega')+\omega_0/c(\omega_0)-(\omega'+\omega_0)(d/d\omega)(\omega/c(\omega))\big|_{\omega=\omega_0}}{(\omega'+\omega_0)^2}-\mathrm{Im}\,\Lambda_2(\omega,\omega_0)\right)\frac{d\omega'}{\omega'+\omega} \tag{A6}$$

and

$$\frac{\omega}{c(\omega)}=\frac{\omega_0}{c(\omega_0)}+(\omega-\omega_0)\frac{d}{d\omega}\frac{\omega}{c(\omega)}\bigg|_{\omega=\omega_0}$$
$$-\frac{(\omega-\omega_0)^2}{\pi}\int_0^\infty\frac{\mathrm{Re}\,\Lambda_2(\omega',\omega_0)-\mathrm{Re}\,\Lambda_2(\omega,\omega_0)}{\omega'-\omega}$$
$$\times d\omega'+\frac{(\omega-\omega_0)^2}{\pi}$$
$$\times\int_0^\infty\left(\frac{\alpha(\omega')+\alpha(\omega_0)-(\omega'+\omega_0)\alpha'(\omega_0)}{(\omega'+\omega_0)^2}\right.$$
$$\left.+\mathrm{Re}\,\Lambda_2(\omega,\omega_0)\right)\frac{d\omega'}{\omega'+\omega}, \tag{A7}$$

where

$$\mathrm{Im}\,\Lambda_2(\omega,\omega_0)=\frac{\omega/c(\omega)-\omega_0/c(\omega_0)-(\omega-\omega_0)(d/d\omega)\omega/c(\omega)\big|_{\omega=\omega_0}}{(\omega-\omega_0)^2}, \tag{A8}$$

and

$$\mathrm{Re}\,\Lambda_2(\omega,\omega_0)=\frac{-\alpha(\omega)+\alpha(\omega_0)+(\omega-\omega_0)(d/d\omega)\alpha(\omega)\big|_{\omega=\omega_0}}{(\omega-\omega_0)^2}. \tag{A9}$$

When using the truncated relations, the inclusion of the subtraction constants $\mathrm{Im}\,\Lambda_2(\omega,\omega_0)$ and $\mathrm{Re}\,\Lambda_2(\omega,\omega_0)$ can have a nonzero contribution to the integrals but will be important to eliminate one of the artifacts as discussed in Sec. II C 1. (The above expressions can be related to more commonly quoted forms of the K–K relations by taking the appropriate limits, such as $\omega_0\to 0$.)

3. Expansion form of the K–K relations

An alternate way to express the above relations are given by the ''expansion'' forms,

$$\mathrm{Re}\,\gamma(\omega)-\mathrm{Re}\,\gamma(\omega_0)-(\omega-\omega_0)\frac{d}{d\omega}\mathrm{Re}\,\gamma(\omega)\big|_{\omega=\omega_0}$$
$$=\lim_{|\Omega|\to\infty}\left(\frac{1}{\pi}\int_{-\Omega}^{\Omega}\frac{\mathrm{Im}\,\gamma(\omega')-\mathrm{Im}\,\gamma(\omega)}{\omega'-\omega}d\omega'\right.$$
$$-\frac{1}{\pi}\int_{-\Omega}^{\Omega}\frac{\mathrm{Im}\,\gamma(\omega')-\mathrm{Im}\,\gamma(\omega_0)}{\omega'-\omega_0}d\omega'-(\omega-\omega_0)$$
$$\left.\times\frac{1}{\pi}\int_{-\Omega}^{\Omega}\mathrm{Im}\,\Lambda_2(\omega',\omega_0)d\omega'\right) \tag{A10}$$

and

$$\mathrm{Im}\,\gamma(\omega)-\mathrm{Im}\,\gamma(\omega_0)-(\omega-\omega_0)\frac{d}{d\omega}\mathrm{Im}\,\gamma(\omega)\big|_{\omega=\omega_0}$$
$$=\lim_{|\Omega|\to\infty}\left(-\frac{1}{\pi}\int_{-\Omega}^{\Omega}\frac{\mathrm{Re}\,\gamma(\omega')-\mathrm{Re}\,\gamma(\omega)}{\omega'-\omega}d\omega'\right.$$
$$+\frac{1}{\pi}\int_{-\Omega}^{\Omega}\frac{\mathrm{Re}\,\gamma(\omega')-\mathrm{Re}\,\gamma(\omega_0)}{\omega'-\omega_0}d\omega'+(\omega-\omega_0)$$
$$\left.\times\frac{1}{\pi}\int_{-\Omega}^{\Omega}\mathrm{Re}\,\Lambda_2(\omega',\omega_0)d\omega'\right). \tag{A11}$$

The meaning of ''twice-subtracted'' is clear in these forms—the two last terms on the right-hand sides can be necessary to subtract away any divergences that may arise. If there is some high frequency above which $\mathrm{Im}\,\gamma(\omega)\sim\mathrm{sgn}(\omega)|\omega|^r$, r

$\leqslant 1$, the first two integrals in Eq. (A10) diverge linearly at worst. Taken together in the $|\Omega| \to \infty$ limit, the divergences of these two cancel out. The third integral in (A10) is always convergent for this case. If Re $\gamma(\omega) \sim |\omega|^r$, $r \leqslant 1$, in the high-frequency limit, the only divergence in (A11) is a logarithmic one which arises only when $r=1$. In this case all three integrals must work together to eliminate the infinity. The first terms on both sides of either relation are the Hilbert transform relations.

The terms beyond the Hilbert transform can be thought of as a power series expansion to first order in $\omega - \omega_0$. Going to a higher-order relation is equivalent to adding higher-order terms in the series, hence the label ''expansion'' for these forms of the relations. Thus removing the linear terms on both sides of Eqs. (A10) and (A11) leaves the once-subtracted relations. This ''expansion'' picture is useful when examining finite interval problems with systems exhibiting prominent local variations like the resonances in the Albunex data.

If $\gamma(\omega)$ obeyed the Hilbert transforms alone, then each member on the left sides of Eqs. (A10) and (A11) would be equal term-by-term to an integral factor on the right. If the minimum number of subtractions required for convergence were one, then only the last terms on each side would be a strict equivalence to one another. With two subtractions required, there is no term-by-term correspondence since each integral taken alone may be divergent over the infinite interval.

APPENDIX B: TWICE-SUBTRACTED BANDLIMITED K–K RELATIONS APPLIED TO A HILBERT TRANSFORM PAIR

Here we calculate a closed form expression for the twice-subtracted finite bandwidth K–K relations applied to the Hilbert transform pair,

$$f(\omega) = \frac{1}{1+(\omega-\omega_r)^2/\Gamma^2} + \frac{1}{1+(\omega+\omega_r)^2/\Gamma^2}, \tag{B1}$$

and

$$g(\omega) = \frac{(\omega-\omega_r)/\Gamma}{1+(\omega-\omega_r)^2/\Gamma^2} + \frac{(\omega+\omega_r)/\Gamma}{1+(\omega+\omega_r)^2/\Gamma^2}. \tag{B2}$$

These functions are the real and imaginary components of the Fourier transform $H(\omega) = (1/\Gamma\sqrt{2\pi})(f(\omega) + ig(\omega))$ of a damped harmonic oscillator $h(t) = 2\vartheta(t)\exp(-\Gamma t)\cos(\omega_r t)$, where $\vartheta(t)$ is the unit step function. These pairs of curves are also known as the in-phase and quadrature components of a Lorentzian spectral line.

Following the steps described in Appendix A for the twice-subtracted dispersion relations, we form the second-order difference functions for the pair:

$$\frac{F_2^{(0,\infty)}(\omega,\omega_0)}{(\omega-\omega_0)^2} \equiv \mathrm{Re}\,\Lambda_2(\omega,\omega_0) = \frac{f(\omega)-f(\omega_0)-(\omega-\omega_0)f'(\omega_0)}{(\omega-\omega_0)^2}, \tag{B3}$$

and

$$\frac{G_2^{(0,\infty)}(\omega,\omega_0)}{(\omega-\omega_0)^2} \equiv \mathrm{Im}\,\Lambda_2(\omega,\omega_0) = \frac{g(\omega)-g(\omega_0)-(\omega-\omega_0)g'(\omega_0)}{(\omega-\omega_0)^2}. \tag{B4}$$

We apply the twice-subtracted finite-interval relations [the truncated forms of Eqs. (A4) and (A5)] to these functions and then solve for $f(\omega)$ and $g(\omega)$. Using the parity properties of $f(\omega)$, $g(\omega)$, and their derivatives, we calculate the twice-subtracted K–K transforms over positive frequency:

$$\begin{aligned} F_2^{(\omega_a,\omega_b)}(\omega,\omega_0) &= \frac{(\omega-\omega_0)^2}{\pi}\int_{\omega_a}^{\omega_b}\left(\frac{G_2^{(0,\infty)}(\omega',\omega_0)}{(\omega'-\omega_0)^2} - \frac{G_2^{(0,\infty)}(\omega,\omega_0)}{(\omega-\omega_0)^2}\right) \\ &\quad\times\frac{d\omega'}{\omega'-\omega} + \frac{(\omega-\omega_0)^2}{\pi} \\ &\quad\times\int_{\omega_a}^{\omega_b}\left(\frac{g(\omega')+g(\omega_0)-(\omega'+\omega_0)g'(\omega_0)}{(\omega'+\omega_0)^2}\right. \\ &\quad\left.+\frac{G_2^{(0,\infty)}(\omega,\omega_0)}{(\omega-\omega_0)^2}\right)\frac{d\omega'}{\omega'+\omega}, \end{aligned} \tag{B5}$$

and

$$\begin{aligned} G_2^{(\omega_a,\omega_b)}(\omega,\omega_0) &= -\frac{(\omega-\omega_0)^2}{\pi}\int_{\omega_a}^{\omega_b}\left(\frac{F_2^{(0,\infty)}(\omega',\omega_0)}{(\omega'-\omega_0)^2} - \frac{F_2^{(0,\infty)}(\omega,\omega_0)}{(\omega-\omega_0)^2}\right) \\ &\quad\times\frac{d\omega'}{\omega'-\omega} + \frac{(\omega-\omega_0)^2}{\pi} \\ &\quad\times\int_{\omega_a}^{\omega_b}\left(\frac{f(\omega')-f(\omega_0)+(\omega'+\omega_0)f'(\omega_0)}{(\omega'+\omega_0)^2}\right. \\ &\quad\left.-\frac{F_2^{(0,\infty)}(\omega,\omega_0)}{(\omega-\omega_0)^2}\right)\frac{d\omega'}{\omega'+\omega}. \end{aligned} \tag{B6}$$

The resulting expressions are

$$\begin{aligned} F_2^{(\omega_a,\omega_b)}(\omega,\omega_0) &= [f(\omega)-f(\omega_0)-(\omega-\omega_0)f'(\omega_0)] \\ &\quad\times\theta(\omega_a,\omega_b,\omega_r,\Gamma) + [g_s(\omega)-g_s(\omega_0) \\ &\quad-(\omega-\omega_0)g_s'(\omega_0)]\varphi(\omega_a,\omega_b,\omega_r,\Gamma) \end{aligned} \tag{B7}$$

and

$$\begin{aligned} G_2^{(\omega_a,\omega_b)}(\omega,\omega_0) &= [g(\omega)-g(\omega_0)-(\omega-\omega_0)g'(\omega_0)] \\ &\quad\times\theta(\omega_a,\omega_b,\omega_r,\Gamma) - [f_s(\omega)-f_s(\omega_0) \\ &\quad-(\omega-\omega_0)f_s'(\omega_0)]\varphi(\omega_a,\omega_b,\omega_r,\Gamma), \end{aligned} \tag{B8}$$

where

$$\theta(\omega_a,\omega_b,\omega_r,\Gamma)=\frac{1}{\pi}\left(\arctan\frac{\omega_b-\omega_r}{\Gamma}+\arctan\frac{\omega_b+\omega_r}{\Gamma}-\arctan\frac{\omega_a-\omega_r}{\Gamma}-\arctan\frac{\omega_a+\omega_r}{\Gamma}\right), \quad \text{(B9)}$$

$$\varphi(\omega_a,\omega_b,\omega_r,\Gamma)=\frac{1}{2\pi}\ln\left|\frac{\Gamma^2+(\omega_b+\omega_r)^2}{\Gamma^2+(\omega_b-\omega_r)^2}\,\frac{\Gamma^2+(\omega_a-\omega_r)^2}{\Gamma^2+(\omega_a+\omega_r)^2}\right|, \quad \text{(B10)}$$

$$g_s(\omega)=\frac{(\omega-\omega_r)/\Gamma}{1+(\omega-\omega_r)^2/\Gamma^2}-\frac{(\omega+\omega_r)/\Gamma}{1+(\omega+\omega_r)^2/\Gamma^2}, \quad \text{(B11)}$$

and

$$f_s(\omega)=\frac{1}{1+(\omega-\omega_r)^2/\Gamma}-\frac{1}{1+(\omega+\omega_r)^2/\Gamma^2}. \quad \text{(B12)}$$

[Note that the inclusion of the subtraction constants $G_2^{(0,\infty)}(\omega,\omega_0)/(\omega-\omega_0)^2$ and $F_2^{(0,\infty)}(\omega,\omega_0)/(\omega-\omega_0)^2$ in Eqs. (B5) and (B6), respectively, cancel out terms like those discussed in Sec. II C 1 of this paper and also leave no artifacts in the above results.]

Both of the results are similar in form. Each expression [Eqs. (B7) and (B8)] is the sum of two terms, with each term comprised of two distinct factors. The first term in each expression contains the proper result to the exact (i.e., entire spectrum) dispersion relation as the first factor that is in turn multiplied by the artifactual factor $\theta(\omega_a,\omega_b,\omega_r,\Gamma)$. In general $\theta(\omega_a,\omega_b,\omega_r,\Gamma)$ is between 0 and 1 for $\omega_b>\omega_a>0$ and goes to unity when $\omega_a\to 0$ and $\omega_b\to\infty$. Using $\omega_b/2\pi=15$, $\omega_a/2\pi=1$, $\omega_r/2\pi=2.3$, and $\Gamma/2\pi=1$, this factor is 0.84. If the other artifacts are ignored (e.g., $\varphi=0$), the desired result of the calculation is still reduced in scale by 16% for these specific parameters.

The second term of each expression [Eqs. (B7) and (B8)] contains the logarithmic factor $\varphi(\omega_a,\omega_b,\omega_r,\Gamma)$. In the limit of $\omega_b\to\infty$, $\omega_a\to 0$ this factor goes to zero as expected. [$\varphi(\omega_a,\omega_b,\omega_r,\Gamma)$ also has a zero when $\omega_b\omega_a=\Gamma^2+\omega_r^2$.] For $\Gamma/2\pi=1$ MHz and $\omega_r/2\pi=2.3$ MHz, $\varphi(\omega_a,\omega_b,\omega_r,\Gamma)$ is -0.14. Over values of the parameters $\omega_a,\omega_b,\omega_r$ and Γ, relevant for this work, the $f_s(\omega)$ and $g_s(\omega)$ terms that multiply φ are roughly equivalent to $f(\omega)$ and $g(\omega)$, respectively. Over most of the spectrum of interest this purely artifactual second term is smaller in magnitude than the functions we are trying to recover.

The bandlimited estimates for $f(\omega)$ and $g(\omega)$, denoted as $f_2^{(\omega_a,\omega_b)}(\omega)$ and $g_2^{(\omega_a,\omega_b)}(\omega)$, are defined as

$$f_2^{(\omega_a,\omega_b)}(\omega)\equiv F_2^{(\omega_a,\omega_b)}(\omega,\omega_0)+f(\omega_0)+(\omega-\omega_0)f'(\omega_0) \quad \text{(B13)}$$

and

$$g_2^{(\omega_a,\omega_b)}(\omega)\equiv G_2^{(\omega_a,\omega_b)}(\omega,\omega_0)+g(\omega_0)+(\omega-\omega_0)g'(\omega_0). \quad \text{(B14)}$$

Using Eqs. (B7) and (B13) we obtain

$$f_2^{(\omega_a,\omega_b)}(\omega)=f(\omega)\,\theta(\omega_a,\omega_b,\omega_r,\Gamma)+f(\omega_0)(1-\theta(\omega_a,\omega_b,\omega_r,\Gamma))+(g_s(\omega)-g_s(\omega_0))\varphi(\omega_a,\omega_b,\omega_r,\Gamma)+(\omega-\omega_0)\times[f'(\omega_0)(1-\theta(\omega_a,\omega_b,\omega_r,\Gamma))-g_s'(\omega_0)\varphi(\omega_a,\omega_b,\omega_r,\Gamma)]. \quad \text{(B15)}$$

From Eqs. (B8) and (B14),

$$g_2^{(\omega_a,\omega_b)}(\omega)=g(\omega)\,\theta(\omega_a,\omega_b,\omega_r,\Gamma)-(f_s(\omega)-f_s(\omega_0))\varphi(\omega_a,\omega_b,\omega_r,\Gamma)+g(\omega_0)(1-\theta(\omega_a,\omega_b,\omega_r,\Gamma))+(\omega-\omega_0)[g'(\omega_0)(1-\theta(\omega_a,\omega_b,\omega_r,\Gamma))+f_s'(\omega_0)\varphi(\omega_a,\omega_b,\omega_r,\Gamma)]. \quad \text{(B16)}$$

These two results are used to calculate the artifacts in Sec. II C 2.

[1] R. d. L. Kronig, ''On the theory of dispersion of x-rays,'' J. Opt. Soc. Am. **12**(6), 547–557 (1926).
[2] H. A. Kramers, ''La diffusion de la lumiere par les atomes,'' Atti. Congr. Intern. Fisici. Como. **2**, 545 (1927).
[3] J. S. Toll, ''Causality and the dispersion relations: logical foundations,'' Phys. Rev. **104**(6), 1760–1770 (1956).
[4] G. Arfken, *Mathematical Methods for Physicists*, 3rd ed. (Academic, New York, 1985).
[5] R. L. Weaver and Y.-H. Pao, ''Dispersion relations for linear wave propagation in homogeneous and inhomogeneous media,'' J. Math. Phys. **22**, 1909–1918 (1981).
[6] T. S. Robinson, ''Optical constants by reflection,'' Proc. Phys. Soc. London, Sect. B **65**, 910–911 (1952).
[7] T. S. Robinson and W. C. Price, ''The determination of infra-red absorption spectra from reflection measurements,'' Proc. Phys. Soc. London, Sect. B **66**, 969–974 (1953).
[8] H. R. Philipp and E. A. Taft, ''Optical constants of germanium in the region 1 to 10 eV,'' Phys. Rev. **113**(4), 1002–1005 (1959).
[9] V. L. Ginzberg, ''Concerning the general relationship between absorption and dispersion of sound waves,'' Sov. Phys. Acoust. **1**, 32–41 (1955).
[10] C. W. Horton, Sr., ''Dispersion relationships in sediments and sea water,'' J. Acoust. Soc. Am. **55**, 547–549 (1974).
[11] M. O'Donnell, E. T. Jaynes, and J. G. Miller, ''General relationships between ultrasonic attenuation and dispersion,'' J. Acoust. Soc. Am. **63**, 1935 (1978).
[12] K. Aki and P. G. Richards, *Quantitative Seismology: Theory and Methods* (Freeman, San Franscisco, 1980), pp. 171–175.
[13] M. O'Donnell, E. T. Jaynes, and J. G. Miller, ''Kramers–Kronig relationship between ultrasonic attenuation and phase velocity,'' J. Acoust. Soc. Am. **69**, 696–701 (1981).
[14] J. H. Rose, D. K. Hsu, and L. Adler, ''Ultrasonic characterization of porosity using the Kramers–Kronig relations,'' J. Phys. (France) **46**, (suppl. No. 12), C10-787–C10-790 (1985).
[15] D. J. Wingham, ''The dispersion of sound in sediment,'' J. Acoust. Soc. Am. **78**, 1757–1760 (1985).
[16] A. I. Beltzer and N. Brauner, ''Shear waves in polycrystalline media and modifications of the Keller approximation,'' Int. J. Solids Struct. **23**(1), 201–209 (1987).
[17] M. S. Hughes, M. R. Kaufmann, and J. G. Miller, ''Local approximations to the Kramers–Kronig relations obtained using Burmann's theorem,'' (unpublished, 1987).
[18] I. Lerche, ''Inversion of apparent attenuation and phase velocity to determine scattering parameters,'' Geophys. J. **95**, 149–161 (1988).
[19] A. I. Beltzer, *Acoustics of Solids* (Springer-Verlag, New York, 1988).
[20] C. C. Lee, M. Lahham, and B. G. Martin, ''Experimental verification of the Kramers–Kronig relationship for acoustic waves,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **37**(4), 286–294 (1990).
[21] S. Temkin, ''Attenuation and dispersion of sound in bubbly fluids via the Kramers–Kronig relations,'' J. Fluid Mech. **211**, 61–72 (1990).
[22] Y. C. Angel and J. D. Achenbach, ''Attenuation and speed of sound of antiplane waves in a cracked solid using the Kramers–Kronig relations,'' J. Acoust. Soc. Am. **90**, 2757–2762 (1991).
[23] P. H. Brodeur, M. S. Hall, and C. Esworthy, ''Sound dispersion and at-

tenuation in the thickness direction of paper materials,'' J. Acoust. Soc. Am. **94**, 2215–2225 (1993).
[24] B. F. Pouet and N. J. P. Rasolofosaon, ''Measurements of broadband intrinsic ultrasonic attenuation and dispersion in solids with laser techniques,'' J. Acoust. Soc. Am. **93**, 1286–1292 (1993).
[25] D. Zellouf, Y. Jayet, N. Saintpierre, J. Tatibouet, and J. C. Baboux, ''Ultrasonic spectroscopy in polymeric materials—Application of the Kramers–Kronig relations,'' J. Appl. Phys. **80**, 2728–2732 (1996).
[26] E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, 2nd ed. (Oxford U.P., Oxford, 1948).
[27] H. M. Nussenveig, *Causality and Dispersion Relations* (Academic, New York, 1972).
[28] C. W. Horton, Sr., ''Comment on 'Kramers–Kronig relationship between ultrasonic attenuation and phase velocity' [J. Acoust. Soc. Am. **69**, 696–701 (1981)],'' J. Acoust. Soc. Am. **70**, 1182 (1981).
[29] H. C. Booij and G. P. J. M. Thoone, ''Generalization of Kramers–Kronig transforms and some approximations of relations between viscoelastic quantities,'' Rheol. Acta **21**(1), 15–24 (1982).
[30] K. R. Waters, M. S. Hughes, J. Mobley, G. H. Brandenburger, and J. G. Miller, ''On the applicability of Kramers–Kronig relations for media with ultrasonic attenuation obeying a frequency power law,'' J. Acoust. Soc. Am. **108**, 556–563 (2000).
[31] K. R. Waters, M. S. Hughes, J. Mobley, G. H. Brandenburger, and J. G. Miller, ''Kramers–Kronig dispersion relations for ultrasonic attenuation obeying a frequency power law,'' Proceedings of the IEEE Ultrasonics Symposium **1**, 537–541 (1999).
[32] T. L. Szabo, ''Time-domain wave equations for lossy media obeying a frequency power-law,'' J. Acoust. Soc. Am. **96**, 491–500 (1994).
[33] T. L. Szabo, ''Causal theories and data for acoustic attenuation obeying a frequency power law,'' J. Acoust. Soc. Am. **97**, 14–24 (1995).
[34] Z. Ye, ''Acoustic dispersion and attenuation in many spherical scatterer systems and the Kramers–Kronig relations,'' J. Acoust. Soc. Am. **101**, 3299–3305 (1997).
[35] R. Hulthen, ''Kramers–Kronig relations generalized-on dispersion relations for finite frequency intervals—a spectrum restoring filter,'' J. Opt. Soc. Am. **72**(6), 794–803 (1982).
[36] M. E. Orazem, J. M. Esteban, and O. C. Moghissi, ''Practical applications of the Kramers–Kronig relations,'' Corrosion (Houston) **47**(4), 248–259 (1991).
[37] B. A. Boukamp, ''Practical application of the Kramers–Kronig transformation on impedance measurement in solid-state electrochemistry,'' Solid State Ionics **62**(1–2), 131–141 (1993).
[38] G. W. Milton, D. J. Eyre, and J. V. Mantese, ''Finite frequency range Kramers–Kronig relations: bounds on the dispersion,'' Phys. Rev. Lett. **79**(16), 3062–3065 (1997).
[39] H. J. Bowlden and J. K. Wilmshurst, ''Evaluation of the One-Angle Reflection Technique for the Determination of Optical Constants,'' J. Opt. Soc. Am. **53**(9), 1073–1078 (1963).
[40] J. Mobley, J. N. Marsh, C. S. Hall, M. S. Hughes, J. G. Miller, and G. H. Brandenburger, ''Broadband measurements of phase velocity in Albunex suspensions,'' J. Acoust. Soc. Am. **103**, 2145–2153 (1998).
[41] J. Mobley, K. R. Waters, C. S. Hall, J. N. Marsh, M. S. Hughes, G. H. Brandenburger, and J. G. Miller, ''Measurements and predictions of the phase velocity and attenuation coefficient in suspensions of elastic microspheres,'' J. Acoust. Soc. Am. **106**, 652–659 (1999).
[42] J. N. Marsh, C. S. Hall, M. S. Hughes, J. Mobley, J. G. Miller, and G. H. Brandenburger, ''Broadband through-transmission signal loss measurements of Albunex suspensions at concentrations approaching *in vivo* doses,'' J. Acoust. Soc. Am. **101**, 1155–1161 (1997).
[43] R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain* (American Mathematical Society, New York, 1934), pp. 16–17.
[44] E. Pfaffelhuber, ''Generalized impulse response and causality,'' IEEE Trans. Circuit Theory **CT-18**, 218–223 (1971).

Transverse curvature of the acoustic slowness surface in crystal symmetry planes and associated phonon focusing cusps

A. L. Shuvalov[a] and A. G. Every
Department of Physics, University of the Witwatersrand, Johannesburg 2001, South Africa

(Received 3 December 1999; accepted for publication 12 August 2000)

Conditions are derived for the existence of focusing cusps in ballistic phonon intensity patterns for propagation directions in crystal symmetry planes. Line caustics are known to be associated with lines of vanishing Gaussian curvature (parabolic lines) on the acoustic slowness surface, while cusps are associated specifically with points where the direction of vanishing principal curvature is parallel to the parabolic line. A parabolic line meets a crystal symmetry plane σ at a right angle, and so it is the vanishing of the slowness-surface curvature transverse to σ that conditions the existence of a cusp. A relation for the transverse curvature is derived and analyzed. It is shown that in an arbitrary symmetry plane σ there may be up to four pairs of inversion-equivalent cuspidal points for SH (out-of-plane polarized) waves, and up to eight pairs of cuspidal points associated with the in-plane polarized (usually quasi-transverse) waves. In tetragonal crystals, the symmetry planes containing the four-fold axis can have at most two pairs of cusps for the SH waves and up to six pairs of cusps for the in-plane waves. In cubic crystals, the face symmetry planes σ cannot have cuspidal points for SH waves, as is known, while four pairs of cusps for in-plane waves exist in σ if and only if the outer-most slowness sheet has a concave region embracing the four-fold axis. The points of vanishing transverse curvature on the slowness surface in symmetry planes of tetragonal and cubic media are identified by concise relations, facilitating their explicit analysis. [S0001-4966(00)03311-7]

PACS numbers: 43.35.Pt [HEB]

I. INTRODUCTION

The acoustic slowness surface depicts the dependence of the inverse phase velocity of elastic waves upon propagation direction, and is one of the basic tools in elucidating various anisotropy-based phenomena in crystal acoustics and low-temperature phonon physics. The geometry of the slowness surface determines the shape of the ray surface, which represents the wave front of an elastic disturbance radiated from a point source. In consequence, the curvature of the slowness surface is intimately involved in the interpretation and computation of ultrasonic and phonon images, transient-wave propagation, ballistic phonon transport in crystals, and other related effects.[1] The most striking features of phonon images are their caustics arising from intense focusing of acoustic energy. The caustics are associated with parabolic lines (lines of zero Gaussian curvature) on the slowness surface.

The theoretical background describing the slowness and ray surfaces for a given anisotropic medium has been well developed, and extensive experimental and computer-generated data are available at present for different crystal materials, e.g., Refs. 1–8. However, for reasons of algebraic complexity, explicit analytical findings, establishing the curvature of slowness surface and conditioning its sign in terms of elastic coefficients, have been limited to certain particular cases related, for instance, to the symmetry axes[9,10] and the conical-type degeneracy of phase velocities.[11] The shape of the sheets of the slowness surface in a section by a symmetry plane has been studied in a number of papers.[9–18] At the same time, their analysis has been restricted to the in-plane curvature of the slowness surface and leaves important questions unattended. Knowledge of the transverse principal curvature is essential to gain a full picture of the geometrical features of the slowness surface and hence of the possible focusing properties for propagation in a symmetry plane. Vanishing of specifically the transverse curvature of the slowness surface in a symmetry plane, in fact, signals the occurrence of a higher-order singularity (catastrophe) in the focusing caustics pattern, namely, the cuspidal point at which two fold line caustics meet.[2,19,20] Note that folds and cusps of the focusing caustics entail different types of the power-law falloff with distance for monochromatic radiation and of the time dependence for wave-arrival singularities.[19,21]

Motivated by the above-mentioned considerations, the present article is concerned with the explicit derivation and the analysis of the transverse principal curvature of the slowness surface in a section by a symmetry plane. It is hoped that the results obtained, along with the previously established criteria for the in-plane shape of the slowness surface and for its different configuration types near symmetry axes and conical points, will provide further insight into the properties of the slowness and ray surfaces in anisotropic media, and will facilitate predicting of their specific features for a given material.

II. DERIVATION OF THE TRANSVERSE PRINCIPAL CURVATURE

Consider plane wave propagation in an elastic medium of density ρ and elastic stiffness tensor c_{ijkl}. For a given

[a]Permanent address: Institute of Crystallography, Leninskii pr. 59, Moscow 117333, Russia.

wave normal $\mathbf{m}$ ($\mathbf{m}^2=1$) there are three modes with phase velocities v_α and mutually orthogonal polarizations $\mathbf{A}_\alpha$ governed by the Christoffel equation

$$(\mathbf{\Gamma}-\rho v_\alpha^2\mathbf{I})\mathbf{A}_\alpha=0, \quad \alpha=1,2,3, \tag{1}$$

where $\Gamma_{ik}=m_j c_{ijkl} m_l$ is the Christoffel tensor. The wave normal $\mathbf{m}$ can be parametrized through the polar angle θ measured, say, from the axis X_3 of the rectangular coordinate system, and the azimuthal angle φ measured from the axis X_1. The slowness surface, characterizing anisotropy of the inverse phase velocities, consists of three nested centrosymmetric sheets

$$S_\alpha: \mathbf{s}_\alpha(\theta,\varphi)=v_\alpha^{-1}(\theta,\varphi)\mathbf{m}(\theta,\varphi), \quad \alpha=1,2,3, \tag{2}$$

corresponding to the slowest, intermediate, and fastest wave branches. The direction of the ray (energy flux) vector of the αth mode with wave normal $\mathbf{m}(\theta_0,\varphi_0)$ is given by the outward normal $\boldsymbol{\nu}_\alpha(\theta_0,\varphi_0)$ to S_α. The principal curvatures $K_1^{(\alpha)}$, $K_2^{(\alpha)}$ of the slowness surface at a regular point $\mathbf{s}_\alpha(\theta_0,\varphi_0)$ (i.e., away from a point of degeneracy) are the external values of the curvature of the planar curves obtained in sections by the family of planes through $\boldsymbol{\nu}_\alpha(\theta_0,\varphi_0)$.[22] Orientations of the two planes providing those extremal values are said to be the directions of the principal curvatures at $\mathbf{s}_\alpha(\theta_0,\varphi_0)$. The Gaussian curvature is given by $K_{\mathrm{G}}^{(\alpha)}=K_1^{(\alpha)}K_2^{(\alpha)}$. A regular point $\mathbf{s}_\alpha(\theta_0,\varphi_0)$ is said to be elliptical if $K_{\mathrm{G}}^{(\alpha)}>0$, corresponding to the surface being convex at $K_1^{(\alpha)}>0$, $K_2^{(\alpha)}>0$ or concave at $K_1^{(\alpha)}<0$, $K_2^{(\alpha)}<0$; hyperbolic if $K_{\mathrm{G}}^{(\alpha)}<0$; and parabolic if $K_{\mathrm{G}}^{(\alpha)}=0$. The inner of the three sheets S_α is always convex,[23] while the other two sheets may contain domains of the concave and hyperbolic types. They are bounded by lines of parabolic points, which map on the cuspidal edges of the wave surface and focusing line caustics. A point at which the parabolic line is tangent to the direction of vanishing principal curvature maps onto a cuspidal point of the ray surface and a focusing cusp caustic.[2,19,20]

Directions $\mathbf{m}$, in which two slowness sheets S_α come into contact and there is phase-velocity degeneracy, are termed acoustic axes. The tangential-type contact along fourfold axes, and the conical-type contact, which occurs along all threefold axes and also in some nonsymmetric direction, bring about irregular points of the degenerate sheets S_α.[5] Such a point cannot be attributed a specific Gaussian curvature, and it may be the point of intersection of several parabolic lines (see Refs. 10 and 11).

Consider a symmetry plane σ of an orthorhombic medium. Let us define the coordinate system so that σ is the plane X_1X_2, that is, $\theta=\pi/2$, and hence an orientation of the wave normal $\mathbf{m}$ in σ is described by the azimuthal angle φ measured from the axis X_1 parallel to the twofold axis. (The forthcoming derivations are equally valid for the symmetry plane in a monoclinic medium provided that the orientation of the X_1, X_2 axes is chosen which renders $c_{45}=0$.) Addressing the transverse curvature of the slowness surface in the section by the plane σ will require appeal to the approximate form of Christoffel tensor in the close neighborhood of σ up to the quadratic terms in $\Delta\theta=\pi/2-\theta\ll 1$,

$$\mathbf{\Gamma}\approx\begin{pmatrix} \gamma_{11}+(c_{55}-\gamma_{11})(\Delta\theta)^2 & \gamma_{12}[1-(\Delta\theta)^2] & \gamma_{13}(\Delta\theta) \\ \gamma_{12}[1-(\Delta\theta)^2] & \gamma_{22}+(c_{44}-\gamma_{22})(\Delta\theta)^2 & \gamma_{23}(\Delta\theta) \\ \gamma_{13}(\Delta\theta) & \gamma_{23}(\Delta\theta) & \gamma_{33}+(c_{33}-\gamma_{33})(\Delta\theta)^2 \end{pmatrix}, \tag{3}$$

where

$$\begin{aligned} &\gamma_{11}=c_{11}-\Delta_{16}\sin^2\varphi, \quad \gamma_{22}=c_{66}+\Delta_{26}\sin^2\varphi, \\ &\gamma_{33}=c_{55}+\Delta_{45}\sin^2\varphi, \quad \gamma_{12}=d_{12}\sin\varphi\cos\varphi, \\ &\gamma_{13}=d_{13}\cos\varphi, \quad \gamma_{23}=d_{23}\sin\varphi. \end{aligned} \tag{4}$$

Here the elastic coefficients are written in Voigt's notation. We also make use of the quantities

$$\begin{aligned} &\Delta_{ab}\equiv c_{aa}-c_{bb}, \quad a,b=1,...,6; \\ &d_{ij}=c_{iijj}+c_{ijij}, \quad i,j=1,2,3, \quad i\neq j. \end{aligned} \tag{5}$$

According to the Christoffel equation (1), one of the three wave branches with a wave normal in the symmetry plane $\sigma(\theta=\pi/2)$ is polarized orthogonally to σ (the pure transverse or the SH branch), while the polarizations of the other two branches lie in σ. We will hereafter assign the superscript $\alpha=t$ to the parameters of the SH branch, and the superscripts $\alpha=s, f$ to the parameters of the in-plane slow and fast branches, respectively. The phase velocities of the SH and the in-plane waves are given by

$$\begin{aligned} &\rho v_t^2(\varphi)=c_{55}+\Delta_{45}\sin^2\varphi; \\ &\rho v_{s,f}^2(\varphi)=\tfrac{1}{2}[c_{11}+c_{66}-\Delta_{12}\sin^2\varphi \\ &\qquad \mp\sqrt{(\Delta_{16}\cos^2\varphi-\Delta_{26}\sin^2\varphi)^2+d_{12}^2\sin^2 2\varphi}]. \end{aligned} \tag{6}$$

The argument $\theta=\pi/2$ here and in the following derivations is suppressed.

Denote the curves, lying on the sheets of slowness surface in the section by the symmetry plane σ, as $S_\alpha(\varphi)$, $\alpha=t,s,f$. On symmetry grounds, $(v_\alpha)'_\theta=0$ at $\theta=\pi/2$ (hereafter prime denotes differentiation with respect to the subscripted parameter). Hence, the unit normals $\boldsymbol{\nu}_\alpha(\varphi)$ all lie in the mirror plane σ and have the form (see Refs. 16 and 17)

$$\boldsymbol{\nu}_\alpha(\varphi)=\frac{1}{\sqrt{v_\alpha^2+(v_\alpha)'^2_\varphi}}[v_\alpha\mathbf{m}+(v_\alpha)'_\varphi\mathbf{m}'_\varphi]. \tag{7}$$

Due to the symmetry, one of the principal curvatures of the slowness surface at points in the plane σ is the in-plane one,

$$K_1^{(\alpha)}(\varphi)=-\frac{\boldsymbol{\nu}_\alpha\cdot(\mathbf{s}_\alpha)''_{\varphi\varphi}}{|(\mathbf{s}_\alpha)'_\varphi|}=\frac{v_\alpha^4}{[v_\alpha^2+(v_\alpha)'^2_\varphi]^{3/2}}\left[1+\frac{(v_\alpha)''_{\varphi\varphi}}{v_\alpha}\right], \quad (8)$$

which is the curvature for the planar curve $S_\alpha(\varphi)$, and the second is the transverse curvature,

$$K_2^{(\alpha)}(\varphi)=-\frac{\boldsymbol{\nu}_\alpha\cdot(\mathbf{s}_\alpha)''_{\theta\theta}}{|(\mathbf{s}_\alpha)'_\theta|}=\frac{v_\alpha^2}{\sqrt{v_\alpha^2+(v_\alpha)'^2_\varphi}}\left[1+\frac{(v_\alpha^2)''_{\theta\theta}}{2v_\alpha^2}\right], \quad (9)$$

which is the curvature for the planar curve in the section by the transverse plane $\varphi=$ const. For brevity, we will be calling $K_2^{(\alpha)}(\varphi)$ ''the transverse curvature of $S_\alpha(\varphi)$.'' Again because of symmetry, any parabolic line, meeting σ at a regular point, does so at a right angle. It follows that at this point of intersection the vanishing curvature is either normal or parallel to the parabolic line if, respectively, it is the in-plane or the transverse curvature that vanishes (we exclude the singular event of both vanishing). In the former case the intersection point maps onto a fold line of the ray surface, while in the latter case the intersection point maps onto a cuspidal point of the ray surface.[2]

The curvatures $K_1^{(t)}(\varphi)$ and $K_1^{(f)}(\varphi)$ are strictly positive. The properties of the curvature $K_1^{(s)}(\varphi)$ for the slow in-plane branch have been comprehensively sorted out in Refs. 9 and 12–18. On finding the eigenvalues $\rho v_\alpha^2(\theta,\varphi)$ of $\mathbf{\Gamma}$, Eq. (3), for small deviations of θ from $\pi/2$ with the aid of perturbation theory, and then inserting the result into Eq. (9), we obtain the expression

$$\begin{aligned}K_2^{(t)}(\varphi)=&\frac{1}{\rho^3\sqrt{v_t^2+(v_t)'^2_\varphi}(v_s^2-v_t^2)(v_f^2-v_t^2)}\{\{c_{33}[d_{12}^2-(\Delta_{15}+\Delta_{46})(\Delta_{24}+\Delta_{56})]+d_{13}^2(\Delta_{24}+\Delta_{56})+d_{23}^2(\Delta_{15}+\Delta_{46})\\&-2d_{12}d_{13}d_{23}\}\sin^4\varphi-[c_{33}(d_{12}^2-\Delta_{15}\Delta_{24}-\Delta_{46}\Delta_{56}-2\Delta_{15}\Delta_{56})\\&-2d_{12}d_{13}d_{23}+d_{13}^2(\Delta_{24}+2\Delta_{56})+d_{23}^2\Delta_{15}]\sin^2\varphi-\Delta_{56}(c_{33}\Delta_{15}-d_{13}^2)\}\end{aligned} \quad (10)$$

for transverse curvature of the SH branch, and

$$\begin{aligned}K_2^{(\alpha)}(\varphi)=&\frac{1}{\rho^3\sqrt{v_\alpha^2+(v_\alpha)'^2_\varphi}(v_t^2-v_\alpha^2)(v_\beta^2-v_\alpha^2)}\{(c_{55}+c_{44})\\&\times(\rho v_\alpha^2)^2-[c_{55}(\gamma_{22}+\gamma_{33})+c_{44}(\gamma_{11}+\gamma_{33})-\gamma_{13}^2\\&-\gamma_{23}^2]\rho v_\alpha^2+\gamma_{33}(c_{55}\gamma_{22}+c_{44}\gamma_{11})+2\gamma_{12}\gamma_{13}\gamma_{23}\\&-\gamma_{23}^2\gamma_{11}-\gamma_{13}^2\gamma_{22}\},\quad \alpha,\beta=s,f,\end{aligned} \quad (11)$$

for transverse curvature of the in-plane branches [the phase velocities, which appear in (10) and (11), are given by (6)]. Evaluating (10) and (11), in particular at $\cos\varphi=1$ and at $\cos\varphi=0$, leads to the known values of the transverse curvature $K_2^{(\alpha)}(0)$, $K_2^{(\alpha)}(\pi/2)$ at the twofold axes X_1, X_2, respectively [see Eq. (22) in Ref. 10].

For a symmetry plane containing a sixfold axis **e** (zonal planes in hexagonal media), it has been shown that the transverse curvature of the curve $S_\alpha(\varphi)$ vanishes under the condition $[v_\alpha^{-1}(\varphi)\cos\varphi]'_\varphi=0$, which holds if and only if there is the in-plane concavity around **e**, that is, at $c_{11}<d_{13}^2/\Delta_{34}$.[14] All points on the circular parabolic line pertaining to the vanishing of the transverse curvature map onto a single ray parallel to **e**, which is related to the external refraction phenomenon.[9] It results in especially intense phonon focusing along **e**.[24] The case of hexagonal media is excluded from the subsequent discussion.

The transverse curvature is undefined at points of phase-velocity degeneracy, either stipulated by the appropriate symmetry axis in σ or possibly occurring along a nonsymmetric direction. Barring special coincidences among the elastic coefficients, the degeneracy arising in σ involves the SH and one of the in-plane branches (usually, the slow quasi-transverse one). By (6), the condition $v_\alpha(\varphi)=v_t(\varphi)$, $\alpha=s$ or f, gives the equation (e.g., Ref. 25)

$$\begin{aligned}&[d_{12}^2-(\Delta_{15}+\Delta_{46})(\Delta_{24}+\Delta_{56})]\sin^4\varphi\\&\quad-(d_{12}^2-\Delta_{15}\Delta_{24}-\Delta_{46}\Delta_{56}-2\Delta_{15}\Delta_{56})\sin^2\varphi\\&\quad-\Delta_{15}\Delta_{56}=0,\end{aligned} \quad (12)$$

which determines up to four pairs of inversion-equivalent directions of degeneracy $\mathbf{m}_d$ (the acoustic axes) possible in σ. It is essential to note that when the wave normal **m**, varying in σ, passes through a direction $\mathbf{m}_d$ of conical degeneracy between two nested (simply connected) slowness sheets, the SH branch skips from the outer sheet to the inner one, and the in-plane branch involved in the degeneracy (say, $\alpha=s$) does the opposite.[26,27] In other words, the slowness curves $S_t(\varphi)$ and $S_s(\varphi)$, distinguished in σ according to their polarization continuity (SH and in-plane, respectively), are smooth curves intersecting at the conical-degeneracy point, at which each of them moves from one of the nested slowness sheets to the other. As a result, the transverse curvature of $S_t(\varphi)$ and $S_s(\varphi)$ may change sign not only via zero, to give rise to a cuspidal point, but also via infinity at the point of conical degeneracy. If **m**, varying in σ, passes a direction $\mathbf{m}_d$ of tangential degeneracy, then the SH and the in-plane waves remain attached to the same simply connected slowness sheet each, and hence so do the curves $S_t(\varphi)$ and $S_s(\varphi)$.[26,27]

III. CUSPS FOR THE SH BRANCH

A. An arbitrary symmetry plane

The slowness surface at a point $\mathbf{s}_t(\varphi)$, related in a symmetry plane σ to the SH branch, may be either convex or hyperbolic, depending on the sign of the transverse curvature $K_2^{(t)}(\varphi)$ [since $K_1^{(t)}(\varphi)>0$]. When the curve $S_t(\varphi)$ passes through domains of both types, they are separated either by a parabolic point, which of necessity maps onto a focusing cusp, or by a conical point, at which $K_2^{(t)}(\varphi)$ changes sign via infinity. In the general case of a symmetry plane in a monoclinic crystal, the condition

$$K_2^{(t)}(\varphi)=0 \tag{13}$$

is a polynomial equation of the fourth order in $x=\tan\varphi$ (we omit its rather cumbersome explicit form). Hence, for the SH branch in any symmetry plane σ there may be at most four pairs of inversion-equivalent cuspidal points, whose locations follow from (7) for the wave normal directions given by the real roots of (13). Equation (13) remains a full fourth-order equation in $\tan\varphi$ for a symmetry plane σ in a trigonal crystal. According to (10), in the case of a symmetry plane σ in an orthorhombic medium, (13) becomes a quadratic equation in $\sin^2\varphi$, and there are either no cuspidal points or four pairs of them depending on the properties of roots of (13). Once it has been confirmed that only a single cuspidal point is allowed in between the two-fold axes $\varphi=0$ and $\varphi=\pi/2$ in σ, its presence may be established by comparing of signs of the values $K_2^{(t)}(0)$, $K_2^{(t)}(\pi/2)$ at the two-fold axes X_1, X_2,[10]

$$K_2^{(t)}(0)=\frac{1}{\sqrt{\rho c_{55}}}\left(c_{33}-\frac{d_{13}^2}{\Delta_{15}}\right),$$
$$K_2^{(t)}\left(\frac{\pi}{2}\right)=\frac{1}{\sqrt{\rho c_{44}}}\left(c_{33}-\frac{d_{23}^2}{\Delta_{24}}\right) \tag{14}$$

jointly with the criterion for the existence of conical points on $S_t(\varphi)$ in σ [see (12)].

B. Symmetry planes in a tetragonal crystal

1. The symmetry plane normal to the four-fold axis (the X_1X_2 plane)

In this case $S_t(\varphi)$ is a circle, and in consequence the wave normal $\mathbf{m}$ and the ray $\mathbf{v}_t$ are parallel. Nevertheless, the transverse curvature of $S_t(\varphi)$ depends on φ. Equation (13) simplifies to the form

$$\sin^2 2\varphi=\frac{4\Delta_{46}(c_{33}\Delta_{14}-d_{13}^2)}{(\Delta_{16}-d_{12})[c_{33}(c_{11}+c_{12})-2d_{13}^2]}. \tag{15}$$

If the right-hand side of (15) is positive and less than 1, it defines the location of four pairs of inversion-equivalent cuspidal points in σ. The existence of the cuspidal points may be checked by comparing the signs of the transverse curvature $K_2^{(t)}$ at the nonequivalent two-fold axes in σ,[10]

$$K_2^{(1)}(0)=\frac{1}{\sqrt{\rho c_{44}}}\left(c_{33}-\frac{d_{13}^2}{\Delta_{14}}\right),$$
$$K_2^{(t)}\left(\frac{\pi}{2}\right)=\frac{1}{\sqrt{\rho c_{44}}}\left[c_{33}-\frac{2d_{13}^2}{c_{11}+c_{12}-2\Delta_{46}}\right], \tag{16}$$

and considering the availability of the conical points on $S_t(\varphi)$ in σ, which are determined according to (12) by the equation

$$\sin^2 2\varphi=\frac{4\Delta_{46}\Delta_{14}}{\Delta_{16}^2-d_{12}^2}. \tag{17}$$

2. The symmetry plane containing the four-fold axis

Consider, say, one of the symmetry-equivalent planes X_1X_3 or X_2X_3 (the azimuthal angle φ is measured from X_1 or X_2, respectively). At $\sin\varphi=1$, as a result of the tangential-type degeneracy along the four-fold axis, both the numerator and denominator of $K_2^{(t)}$, Eq. (10), turn to zero, so that the transverse curvature at this point depends on the approaching path, i.e. becomes indefinite (see Refs. 5 and 10). Equation (13) has now one solution, which on the appropriate replacement $c_{55}\to c_{66}$, $c_{13}\to c_{12}$ may be written in the form

$$\sin^2\varphi=\frac{\Delta_{46}(c_{11}\Delta_{16}-d_{12}^2)}{c_{11}[d_{23}^2-\Delta_{16}(\Delta_{34}-\Delta_{46})]+d_{12}^2(\Delta_{34}-\Delta_{46})+d_{13}^2(\Delta_{16}-2d_{12})}. \tag{18}$$

Provided that the right-hand side of (18) is positive and less than 1, it gives two symmetry-related wave normal directions, each of them leading to a pair of inversion-equivalent cuspidal points in the plane σ. The location of these cuspidal points in σ ensue on inserting (18) into (7). The replacement of the elastic coefficients $c_{11}\to\frac{1}{2}(c_{11}+c_{12}+2c_{66})$, $c_{12}\to\frac{1}{2}(c_{11}+c_{12}-2c_{66})$, $c_{66}\to\frac{1}{2}(c_{11}-c_{12})$ in Eq. (18) yields the corresponding equation referred for the diagonal symmetry planes {110}.

C. Symmetry planes in a cubic crystal

1. The symmetry planes orthogonal to the four-fold axes (the face plane)

The transverse curvature of S_t is

$$K_2^{(t)}=\frac{\sqrt{c_{44}}}{\sqrt{\rho}(a+b)}[2(b+1)^2-a(a+b)], \tag{19}$$

where we use the notations

$$a=c_{11}/c_{44}, \quad b=c_{12}/c_{44} \tag{20}$$

introduced in Ref. 2. It is seen that the value $K_2^{(t)}$ in (19) is independent of φ, hence there are no cusps associated with the SH waves in the face planes.[2,10] On varying the elastic coefficients, the transverse curvature $K_2^{(t)}$ changes sign simultaneously in all points of the face planes σ. The condition $K_2^{(t)}=0$ provides the equation of line H in the two-parameter space $\{a, b\}$ obtained in Ref. 2. By Ref. 2, line H stipulates the emergence of saddle shape belts stretching along the face planes σ on the slowness sheet corresponding to the fast quasitransverse branch which, as follows, is SH polarized in σ at this state.

2. The diagonal symmetry plane

Equation (18) takes on the form

$$\sin^2\varphi=\frac{(a+b+2)(b+1)-2(a-1)^2}{(3a+7b+10)(b+1)-2(3a+2b-1)(a-1)}. \tag{21}$$

On varying a, b, the two symmetry-related pairs of cuspidal points arrive from the four-fold axes or from the two-fold axis. The former condition, $\sin\varphi=1$, yields the equation of line E in the plane $\{a, b\}$, and the latter condition, $\sin\varphi=0$, gives the equation of line G in the plane $\{a, b\}$.[2] With reference to Ref. 2, the SH waves correspond to the slow quasitransverse branch near line E, and to the fast one near line G.

IV. CUSPS FOR THE IN-PLANE BRANCHES

A. An arbitrary symmetry plane

The common situation is when the fast in-plane waves are also faster than the SH waves everywhere in a symmetry plane σ, i.e., the slowness curve $S_f(\varphi)$ lies on the innermost one of the three slowness sheets which is always convex.[23] We recall that the smooth curves $S_s(\varphi)$ and $S_t(\varphi)$, corresponding to the slow in-plane and the SH waves in σ, may touch or intersect. In the latter case, each of curves on either side of the degeneracy point belongs to different nested sheets of the slowness surface. Both the in-plane and transverse curvatures $K_1^{(s)}(\varphi)$ and $K_2^{(s)}(\varphi)$ of the curve $S_s(\varphi)$ in σ may be either positive or negative, one of them changing sign at the points where a parabolic line meets σ. Such a point, when associated with vanishing of the in-plane curvature $K_1^{(s)}(\varphi)$, maps onto an ''ordinary'' point on a line caustic. If it is a point where the transverse curvature vanishes, i.e.,

$$K_2^{(s)}(\varphi)=0, \tag{22}$$

then this point maps onto a cuspidal point of the wave surface. In the case of a symmetry plane in a monoclinic, orthorhombic, or trigonal crystal, Eq. (22) may be cast into the form of a polynomial equation of the eighth order in $\tan\varphi$ (of the fourth order in $\sin^2\varphi$ for orthorhombic symmetry). Hence, there can be no more than eight pairs of inversion-equivalent cuspidal points for the slow in-plane branch in σ. In the exceptional case of $S_f(\varphi)$ and $S_t(\varphi)$ intersecting in σ, when the curve $S_f(\varphi)$ is not everywhere the innermost one, the above-mentioned occurrences of cuspidal points may be distributed between two in-plane branches.

A reasonable question is whether the above-mentioned upper bound (eight pairs) for the number of cuspidal points for in-plane waves in a symmetry plane σ can be attained for some medium, that is, whether all roots $x=\tan\varphi$ of (22) can be real. An example confirming the realization of this possibility is discussed in the end of this section.

Provided that σ contains a (single) four-fold axis, which entails tangential-type degeneracy of the phase velocity and hence, a ''0/0''-indefiniteness of $K_2^{(s)}$ [see (11)], Eq. (22) reduces to a third-order equation in $\sin^2\varphi$, so that in this case the number of pairs of cuspidal points for the slow in-plane branch in σ cannot exceed six. An example of a tetragonal body, in which a symmetry plane through the four-fold axis has six pairs of cusps for in-plane waves, is also mentioned at the end of the section. On the other hand, such an option does not seem to exist for diagonal planes of cubic crystals according to the analysis of Ref. 2. An example of four pairs of cuspidal points for in-plane waves in cubic diagonal planes is presented in Ref. 2 (see Fig. 5 therein, which shows two cuspidal points within the symmetry-irreducible 90-degree sector for the in-plane polarized waves; note that the cusp marked as #7 belongs to the SH polarization).

B. The symmetry plane normal to the four-fold axis in a tetragonal crystal

In this case condition (22) becomes a quadratic equation in $\cos^2 2\varphi$, namely,

$$\begin{aligned}&[c_{44}(c_{11}+c_{12})+d_{13}^2]^2(\Delta_{16}-d_{12})^2\cos^4 2\varphi-\{(c_{11}+c_{12})\\&\times[c_{44}(\Delta_{14}-\Delta_{46})+d_{13}^2]^2-2d_{12}(c_{44}d_{12}+d_{13}^2)\\&\times[c_{44}(c_{11}+{}_{12})+d_{13}^2]\}(\Delta_{16}-d_{12})\cos^2 2\varphi+d_{12}^2\\&\times\{(c_{44}d_{12}+d_{13}^2)^2-[c_{44}(\Delta_{14}-\Delta_{46})+d_{13}^2]^2\}=0.\end{aligned} \tag{23}$$

Each of the two roots $x=\cos^2 2\varphi$ of (23), provided it is real and such that $0<x<1$, stipulates the existence of four pairs of cuspidal points in σ (one point within the symmetry-irreducible sector $0<\varphi<\pi/4$). Their location follows from (7). We note that the transverse curvature $K_2^{(s)}$ at the two nonequivalent two-fold axes is positive (for, at the same time it is the in-plane curvature for the SH branches referred to the symmetry planes passing through the four-fold axis),[10]

$$K_2^{(s)}(0)=\frac{c_{44}}{\sqrt{\rho c_{66}}}, \quad K_2^{(s)}\left(\frac{\pi}{2}\right)=\frac{\sqrt{2}c_{44}}{\sqrt{\rho(c_{11}-c_{12})}}. \tag{24}$$

Hence, if there are no conical points on S_s in σ [Eq. (17) has no appropriate solutions], then there are either no or eight

pairs of cuspidal points in σ; four pairs of cusps is not an option. Otherwise, if there are four symmetry-equivalent pairs of conical points on S_s, this of necessity stipulates four pairs of cuspidal points in σ. Note that, by (16) and (17), the occurrence of acoustic axes in σ does not apparently necessitate cuspidal points for the SH branch, thus suggesting the possibility that the parabolic line on the simply-connected slowness sheet embraces two conical points and crosses σ at the points of vanishing transverse curvature, which correspond to the in-plane polarized waves, while the SH waves appear on this sheet in between those conical points.

C. The face symmetry plane in a cubic crystal

One of the roots of Eq. (23) is $\cos^2 2\varphi=1$, corresponding to the ''0/0''-indefinite curvature along the four-fold axes in σ, while the other root written in terms of notations (20) is

$$\cos^2 2\varphi=\frac{(b+1)^2[2(b+1)^2+a+b]}{(b-a+2)[(b+1)^2+a+b]^2}. \tag{25}$$

Once the right-hand side is positive and less than 1, it defines four pairs of cuspidal points in σ. It can be confirmed that on varying elastic coefficients the cuspidal points associated with (25) emerge from the four-fold axis. Let $c_{11}>c_{44}$, which is the case for vast majority of cubic media. Then the condition $\cos^2 2\varphi=1$, applied to (25), is the equation of line D on the $\{a, b\}$ plane obtained in Ref. 2 (therefore, this line corresponds to the in-plane polarization). Cubic materials, whose elastic coefficients occupy the domain to the right of line D, have the slowness sheet for slow quasitransverse waves with a concavity around the four-fold axis (that is fifth pattern in Table II in Ref. 10). The points (25), which give rise to cusps, are the points at which the face planes are intersected by the parabolic line between the axial concavity and the neighboring hyperbolic region (see Ref. 2). It follows that at $c_{11}>c_{44}$ this is the only configuration, which contains cuspidal points in the face symmetry plane σ. In the exceptional case $c_{11}<c_{44}$, the right-hand side of (25) being less than 1 also corresponds to the concavity developing from the four-fold axis on the outmost slowness sheet, which now is nondegenerate along the axis and corresponds to the longitudinal mode. We conclude that, for any relation between c_{11} and c_{44}, the concavity of the outmost slowness surface near the four-fold axis is the necessary and sufficient condition for the existence of the cuspidal points in the cube face plane σ. Examples of real and model cubic materials with the appropriate slowness surface may be found in Refs. 2–4.

D. Examples

Consider a cubic crystal with $c_{11}>c_{44}$, whose outmost slowness sheet, S_s, has the concave regions embracing the four-fold axes and hence four pairs of inversion-equivalent cuspidal points in face symmetry planes (one cusp in each symmetry-irreducible sector $0<\varphi<\pi/4$) for in-plane polarized waves, as described above. An example of such a crystal is copper [see Fig. 1(a)]. Application of a tetragonal pertur-

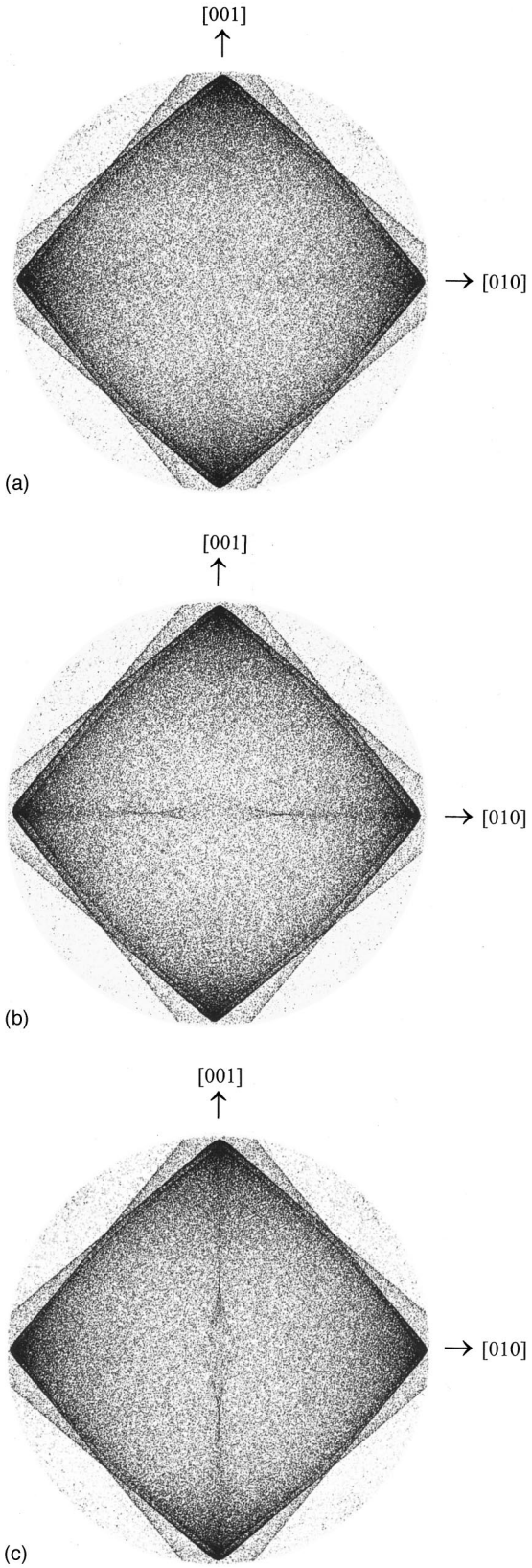


FIG. 1. (a) Focusing pattern for the slow quasitransverse modes of copper in an angular region extending out to 27 degrees around the [100] direction (there are four symmetry cusps in [001] and [010] directions abutting on line caustics to form slightly unfolded hyperbolic umbolic caustics, see Ref. 2). (b) Same as (a), except that c_{66} has been *reduced* by 1%. Same as (a) but with c_{66} *increased* by 1%.

bation $0<(c_{44}-c_{66})/c_{44}\ll 1$ lifts the four-fold rotation symmetry attached to the axes X_1, X_2 and hence splits the tangential degeneracies along those axes into a pair of conical degeneracies each, lying in the planes X_1X_3 and X_2X_3, respectively. Provided that the difference between c_{44} and c_{66} is sufficiently small, the outer slowness sheet acquires a new closed parabolic line around each of the axes X_1, X_2. Each of the lines has the form of two arches, connecting conical points and separating the hyperbolic region, which emerges around the axis, from the surrounding concave region. The slowness curve $S_s(\varphi)$, associated with in-plane polarized waves in the plane X_1X_2, belongs to the outer slowness sheet. The transversal curvature of $S_s(\varphi)$ turns from positive to negative value at the points, where the new-borne parabolic lines cross the plane X_1X_2. This gives rise to two new cuspidal points in this plane near each axis [Fig. 1(b)], so that altogether there are now eight pairs of inversion-equivalent cusps in the symmetry plane X_1X_2 for the in-plane polarized waves. We have therefore confirmed that the upper bound, provided by the degree of Eq. (22) for the number of cuspidal points associated in a symmetry plane, is attainable.

Suppose, next, that the tetragonal perturbation has opposite sign, i.e., $0<(c_{66}-c_{44})/c_{44}\ll 1$. Then the initial tangential degeneracies along the axes X_1 and X_2 split into the pairs of conical degeneracies lying in the plane X_1X_2. In this case, the new-borne closed parabolic lines, connecting the pairwise conical points on the outer slowness sheet, produce two pairs of new inversion equivalent cusps for in-plane polarized waves in each of the symmetry planes X_1X_3 and X_2X_3 [Fig. 1(c)]. Bearing in mind the other four pairs of cusps, initially existing for those wave branches in each of the aforementioned planes, we arrive at the conclusion that here is an example of three pairs of cuspidal points, associated with the in-plane waves in a symmetry plane containing a (single) four-fold axis.

V. SUMMARY

It is known that vanishing of the transverse curvature of the acoustic slowness surface in a section by a symmetry plane σ entails a cuspidal point on the ray surface and associated cusp caustic in the phonon focusing pattern. Analysis of the expression for the transverse curvature yields the conclusion that there may be up to four inversion-equivalent pairs of cuspidal points for the SH branch, and up to eight pairs of cuspidal points for the in-plane polarized waves propagating in an arbitrary symmetry plane. Relatively concise explicit relations identifying location of the points of vanishing transverse curvature are obtained for tetragonal and cubic media. For the symmetry plane σ normal to the four-fold axis in a tetragonal crystal it is shown that the in-plane polarized waves can provide either no or eight pairs of cuspidal points in the absence of conical degeneracy in σ; otherwise four symmetry-equivalent acoustic axes of necessity stipulate four pairs of cuspidal points in σ. The symmetry plane, which passes through the four-fold axis in a tetragonal crystal, can have at most two pairs of cusps for SH waves and up to six pairs of cusps for in-plane waves. In a cubic crystal, the face symmetry plane does not admit cuspidal points for SH waves, as is known; here it is shown that there may be four symmetry-equivalent pairs of cuspidal points for the in-plane polarized waves, arising in those planes if and only if the outmost slowness sheet has a concave region embracing the four-fold axes.

ACKNOWLEDGMENTS

Financial support for this project has been provided by the South African National Research Foundation, and the Russian Foundation for Fundamental Research (Grant No. 97-02-16338).

[1] J. P. Wolfe, *Imaging Phonons* (Cambridge U.P., Cambridge, 1998).
[2] A. G. Every, Phys. Rev. B **24**, 3456 (1981).
[3] A. G. Every and A. J. Stoddart, Phys. Rev. B **34**, 1319 (1985).
[4] D. C. Hurley and J. P. Wolfe, Phys. Rev. B **32**, 2568 (1985).
[5] H. J. Maris, in *Nonequilibrium Phonons in Nonmetallic Crystals*, edited by W. Eisenmenger and A. A. Kaplyanskii (North-Holland, Amsterdam, 1986), p. 51.
[6] A. G. Every, Phys. Rev. B **34**, 2852 (1986).
[7] A. G. Every, Phys. Rev. B **37**, 9964 (1988).
[8] T. Paszkiewicz and M. Wilczynski, in *Dynamic Properties of Solids*, edited by G. K. Norton and A. A. Maradudin (Elsevier Science B. V., Amsterdam, 1995), p. 257.
[9] M. J. K. Musgrave, *Crystal Acoustics* (Holden-Day, San Francisco, 1970).
[10] A. L. Shuvalov and A. G. Every, Phys. Rev. B **53**, 14906 (1996).
[11] A. L. Shuvalov and A. G. Every, J. Acoust. Soc. Am. **101**, 2381 (1997).
[12] C. G. Winternheimer and A. K. McCurdy, Phys. Rev. B **18**, 6576 (1978).
[13] M. J. P. Musgrave, J. Elast. **9**, 105 (1979).
[14] R. G. Payton, *Elastic Wave Propagation in Transversely Isotropic Media* (Nijhoff, The Hague, 1983).
[15] M. J. P. Musgrave and R. G. Payton, J. Elast. **14**, 269 (1984).
[16] K. Y. Kim, Phys. Rev. B **49**, 3713 (1994).
[17] A. L. Shuvalov, Crystallogr. Rep. **42**, 173 (1997); **43**, 717 (E) (1998).
[18] V. I. Alshits and P. Chadwick, Wave Motion **25**, 347 (1997).
[19] M. V. Berry, Adv. Phys. **25**, 1 (1976).
[20] D. Armbruster and G. Dangelmayr, Z. Phys. B: Condens. Matter **52**, 87 (1983).
[21] A. G. Every and K. Y. Kim, J. Acoust. Soc. Am. **95**, 2505 (1994).
[22] C.-C. Hsiung, *A First Course in Differential Geometry* (Wiley, New York, 1981).
[23] G. F. D. Duff, Philos. Trans. R. Soc. London **252**, 249 (1960).
[24] K. Y. Kim, W. Sachse, and A. G. Every, Phys. Rev. Lett. **70**, 3443 (1993); K. Y. Kim, A. G. Every, and W. Sachse, Int. J. Mod. Phys. B **8**, 2327 (1994).
[25] P. Boulanger and M. Hayes, Proc. R. Soc. London, Ser. A **454**, 2323 (1998).
[26] V. I. Alshits and J. Lothe, Sov. Phys. Crystallogr. **24**, 387 (1979).
[27] A. L. Shuvalov, Proc. R. Soc. London, Ser. A **454**, 2911 (1998).

On a time-domain representation of the Kramers–Krönig dispersion relations

Kendall R. Waters
Laboratory for Ultrasonics, Department of Physics, Washington University, St. Louis, Missouri 63130-4899

Michael S. Hughes and Gary H. Brandenburger
Mallinckrodt, Inc., 675 McDonnell Boulevard, Hazelwood, Missouri 63042

James G. Miller
Laboratory for Ultrasonics, Department of Physics, Washington University, St. Louis, Missouri 63130-4899

(Received 15 September 1999; accepted for publication 12 August 2000)

The development of Kramers–Krönig dispersion relations is typically carried out in the frequency domain. An alternative approach known as the time-causal theory develops dispersion relations for media with attenuation obeying a frequency power law through analysis in the time domain [T. L. Szabo, J. Acoust. Soc. Am. **96**, 491–500 (1994)]. Although both approaches predict identical dispersion relations, *it is perceived* that these two approaches are distinct from each other. It is shown, however, that the time-causal theory is in essence a time-domain formulation of the Kramers–Krönig dispersion relations for the special case of media with attenuation obeying a frequency power law. Additionally, it is shown that time-domain representations of the Kramers–Krönig dispersion relations are available for a broader class of media than simply those with power law attenuation. The time-causal theory and the Kramers–Krönig dispersion relations can be viewed as two complementary, yet equivalent, approaches to the study of dispersion. © *2000 Acoustical Society of America.* [S0001-4966(00)03511-6]

PACS numbers: 43.35.Ty [HEB]

I. INTRODUCTION

Ultrasonic characterization of materials is often achieved by measuring the attenuation coefficient and phase velocity. In some experimental situations one ultrasonic parameter may be more accessible than another; for example, measurements of the attenuation coefficient may be more robust than measurements of phase velocity. In these situations, application of the Kramers–Krönig dispersion relations may permit the determination of the phase velocity from the measured attenuation coefficient. Generally speaking, the Kramers–Krönig dispersion relations relate the real and imaginary parts of some susceptibility function. For example, the complex wave number may be chosen as the susceptibility function. The validity of the Kramers–Krönig dispersion relations hinges on the notion of causality. Consequently, the Kramers–Krönig dispersion relations can serve as a check on the causal consistency of a theoretical model.

In the recent literature[1] doubt has been raised regarding the causal validity of the Kramers–Krönig dispersion relations for certain classes of materials. Specifically, media that exhibit attenuation obeying a frequency power law,

$$\alpha(\omega)=\alpha_0|\omega|^y, \tag{1}$$

where α_0 and y are material-dependent, real constants, are claimed to render the Kramers–Krönig dispersion relations invalid for the case $y\geqslant 1$. This view *appears* to be supported by the Paley–Wiener theorem,[2,3] which places limits on the growth of the attenuation at arbitrarily large frequencies, suggesting that $y<1$ is necessary. However, the technique known as the method of subtractions[4] can be employed to develop causally valid dispersion relations.[5] In addition, through use of the method of subtractions the Paley–Wiener theorem is satisfied.

The (initial) *apparent* failing of the Kramers–Krönig dispersion relations for these media with attenuation obeying a frequency power law led to the development of an alternative time-causal theory that provided an inherently causal model.[1] A key to the development of the time-causal dispersion relations was the use of generalized functions, as opposed to ordinary point functions, to represent physical quantities (e.g., attenuation coefficient and phase velocity). Although generalized functions have been used for dispersion measurements in other areas of physics (e.g., particle physics), they do not appear to have found as much use in ultrasonic dispersion measurements, with the exception of the aforementioned time-causal theory. If the complex wave number is represented as a generalized function, a *generalized* Paley–Wiener theorem[6] is available that permits the relaxation of the restriction on the high-frequency behavior of the attenuation coefficient. Specifically, any frequency power law attenuation represented as a generalized function satisfies the generalized Paley–Wiener theorem. It is the development of the Kramers–Krönig dispersion relations using generalized functions which we now review. We then demonstrate the equivalency of the Kramers–Krönig frequency-domain approach to that of the time-causal approach.

In Sec. II, we review the analysis of generalized functions. In particular we consider derivatives, Fourier transforms, and Hilbert transforms of generalized functions, and contrast this to a similar analysis of ordinary point functions. We follow in Sec. III, with the development of the Kramers–

Krönig dispersion relations using generalized functions, and then consider the special case of attenuation obeying a frequency power law. We also look at the inverse Fourier transform of the complex wave number to consider its time-domain representation. In Sec. IV we review the time-causal approach and its so-called time-domain convolution operators. We close with a discussion comparing these two approaches.

II. REVIEW OF DISTRIBUTIONAL ANALYSIS

We often consider a physical quantity as a function, that is, as a mathematical rule which assigns a number to each numerical value of some independent variable. However, the assumption that a physical quantity must be represented by an ordinary function can be too restrictive for many physical applications.[4] An alternative way of describing a physical quantity is as a functional, that is, as a rule which assigns a number to each function in a set of so-called testing functions.[7] If we wish to determine the properties of a physical system, we operate on the system with some testing mechanism and observe how the system responds to the stimulus. The response is a number $f(\varphi)$ depending on both the system f and the testing mechanism φ.[8]

We use the term generalized functions (or distributions) to refer to a variety of linear, continuous functionals on a space of testing functions. A given class (or space) of generalized functions is always paired with a corresponding class of testing functions. The choice of which space of testing functions and distributions to use depends upon the given physical situation. There are several types of generalized functions available, such as Schwartz distributions, tempered distributions, and ultradistributions.[7,9–12] Later we consider the specific space of tempered distributions and its corresponding testing function space.

The power of distributional analysis rests in large part on the facts that every distribution possesses derivatives of all orders (so that differentiation is a continuous operation in this theory) and that testing functions are sufficiently well behaved at $\pm\infty$. As a consequence, differentiation in a distributional sense commutes with various limiting processes such as integration (including, therefore, Fourier and Hilbert transformations).[7] In contrast, analysis of point functions requires additional conditions in order to permit commutation of operations such as differentiation and integration. It should also be noted that differentiation in a distributional sense is not necessarily the same as differentiation of ordinary point functions. Finally, any physical quantity that can be adequately represented as a function can also be characterized as a generalized function.[7]

A. The space *S* of testing functions of rapid descent

The space S consists of all complex-valued functions $\varphi(t)$ that are infinitely smooth and are such that, as $|t| \to \infty$, they and all their partial derivatives decrease to zero faster than every power of $1/|t|$,[7] hence, the name rapid descent. For our purposes, we consider t as the one-dimensional real variable for time. In the language of Lighthill, the space S consists of ''good functions.'' [13] An example of a testing function of rapid descent is $\varphi(t)=e^{-t^2}$.

B. The space *S′* of tempered distributions

A distribution f is said to be *tempered* if it is a linear, continuous functional on the space S of testing functions of rapid descent. Such distributions are also known as distributions of *slow growth*. The space of all distributions of slow growth is denoted by S', and is considered the *dual* (or *conjugate*) *space of* S.[7] In the language of Lighthill, the space S' consists of ''fairly good functions.'' [13] A convenient example of a tempered distribution is an attenuation obeying a frequency power law [Eq. (1)].

For our present purposes, we restrict ourselves to the space S of testing functions of rapid descent and the space S' of tempered distributions.

C. Select properties of tempered distributions

Before we review some properties of tempered distributions, we first want to mention the notation appropriate for distributional analysis. Our choice of notation is common,[7] although variations can be found.[4] As we mentioned above we call the space of tempered distributions the dual of the space of testing functions of rapid descent. The duality notation we employ to represent the tempered distributions as continuous, linear functionals is given by the inner product $\langle f,\varphi\rangle$. We reiterate that for a distribution to make proper sense, it must always be considered paired with a testing function. If we assume the independent variable to be time t, we may represent $\langle f,\varphi\rangle$ as the integral

$$\langle f,\varphi\rangle=\int_{-\infty}^{\infty} f(t)\varphi(t)\,dt. \tag{2}$$

We now consider three properties of tempered distributions: derivatives, Fourier transforms, and Hilbert transforms. Although we review these properties with tempered distributions in mind, many of these properties are valid for other distribution spaces. We state these properties without proof, but refer the reader elsewhere for further details.[7,9,10,14]

Consider a tempered distribution $f(t)$ and the Dirac delta functional $\delta(t)$. The convolution of a derivative of the Dirac delta functional with a tempered distribution $f(t)$ is given by[9]

$$\langle \delta^{(m)}(t)*f(t),\varphi(t)\rangle=\langle f^{(m)}(t),\varphi(t)\rangle. \tag{3}$$

We see that the mth-order derivative of a distribution is equivalent to the convolution of the distribution with the mth-order derivative of the Dirac delta functional.

Let $f(t)$ and $F(\omega)$ be a Fourier transform pair of tempered distributions. It can be shown that the Fourier transform of a tempered distribution is a tempered distribution.[7] Similarly, the inverse Fourier transform of a tempered distribution is also a tempered distribution. We let F be the Fourier transformation operator defined by

$$\langle \mathrm{F}(f(t)),\varphi(\omega)\rangle = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} d\omega\,\varphi(\omega)\int_{-\infty}^{\infty} f(t)e^{i\omega t}\,dt$$
$$= \langle F(\omega),\varphi(\omega)\rangle, \tag{4}$$

and let F^{-1} be the inverse Fourier transformation operator defined by

$$\langle \mathrm{F}^{-1}(F(\omega)),\varphi(t)\rangle = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dt\,\varphi(t)\int_{-\infty}^{\infty} F(\omega)e^{-i\omega t}\,d\omega$$
$$= \langle f(t),\varphi(t)\rangle, \tag{5}$$

where $\varphi(t)$ and $\varphi(\omega)$ are a Fourier transform pair of testing functions of rapid descent.

The generalized Hilbert transformation operator H on a complex tempered distribution $F(\omega)=F_R(\omega)+iF_I(\omega)$ is defined as[10]

$$\langle F(\omega),\varphi(\omega)\rangle = \left\langle \frac{1}{i}\mathrm{H}(F(\omega)),\varphi(\omega)\right\rangle \tag{6a}$$
$$= \left\langle -\frac{(i\omega)^n}{i\pi}\left[\frac{F(\omega)}{(i\omega)^n} * P\left(\frac{1}{\omega}\right)\right],\varphi(\omega)\right\rangle \tag{6b}$$
$$= \int_{-\infty}^{\infty} d\omega\,\varphi(\omega)\frac{(i\omega)^n}{i\pi}P\int_{-\infty}^{\infty}\frac{F(\omega')}{(i\omega')^n}\frac{d\omega'}{\omega'-\omega}, \tag{6c}$$

where P denotes the Cauchy principal value,[15] n is defined as any integer for which $\mathrm{F}^{-1}(F(\omega))=D^n f(t)$, and $f(t)$ is a tempered distribution that is locally square integrable. Here, D^n represents an nth-order derivative. [Please note the use of a different sign convention in Eq. (6) than that of Eq. (3.16) in Ref. 10.] By considering the real and imaginary parts of $F(\omega)$, Eq. (6) may be considered as a generalized Hilbert transform pair. Separating Eq. (6) into its real and imaginary components, one finds

$$\langle F_R(\omega),\varphi(\omega)\rangle = \left\langle -\frac{(i\omega)^k}{\pi}\left[\frac{F_I(\omega)}{(i\omega)^k} * P\left(\frac{1}{\omega}\right)\right],\varphi(\omega)\right\rangle \tag{7a}$$
$$= \langle \mathrm{H}(F_I(\omega)),\varphi(\omega)\rangle \tag{7b}$$

and

$$\langle F_I(\omega),\varphi(\omega)\rangle = \left\langle +\frac{(i\omega)^k}{\pi}\left[\frac{F_R(\omega)}{(i\omega)^k} * P\left(\frac{1}{\omega}\right)\right],\varphi(\omega)\right\rangle \tag{8a}$$
$$= \langle -\mathrm{H}(F_R(\omega)),\varphi(\omega)\rangle. \tag{8b}$$

We recover the conventional Hilbert transform valid for square-integrable functions when $n=0$. We conclude this section with the action of the inverse Fourier transform on the frequency-domain Hilbert transform, given by[10,14]

$$\langle \mathrm{F}^{-1}[\mathrm{H}(F(\omega))],\varphi(t)\rangle = \langle i\,\mathrm{sgn}(t)\mathrm{F}^{-1}(F(\omega)),\varphi(t)\rangle. \tag{9}$$

III. GENERALIZED FUNCTIONS AND THE KRAMERS–KRÖNIG DISPERSION RELATIONS

The linear propagation of ultrasonic waves through an isotropic, lossy medium (e.g., castor oil) is governed by a (one-dimensional) dissipative wave equation

$$\nabla^2 p(z,t) - \Gamma(t)\frac{\partial p(z,t)}{\partial t} - \frac{1}{c_0^2}\frac{\partial^2 p(z,t)}{\partial t^2} = 0, \tag{10}$$

where ∇^2 is the Laplacian operator, p is the pressure field, Γ accounts for a causal dissipation, c_0 is a constant phase velocity, z represents a spatial coordinate, and t represents time. We assume that the pressure field p is a tempered distribution, and recall that distributional differentiation is equivalent to convolution with a derivative of the Dirac delta functional [see Eq. (3)]. We define the product $\Gamma(t)\delta^{(1)}(t)$ as our loss distribution $L_\gamma(t)$. We then rewrite Eq. (10) as

$$\nabla^2 p(z,t) - L_\gamma(t) * p(z,t) - \frac{1}{c_0^2}\frac{\partial^2 p(z,t)}{\partial t^2} = 0. \tag{11}$$

We interpret our loss distribution $L_\gamma(t)$ as a generalized dissipative term. The reader may now recognize the similarity of Eq. (11) to that of Eq. (2) of Ref. 1. We also remind the reader that the pressure field p is a tempered distribution that acts upon testing functions of rapid descent. Equation (11) then is more appropriately written as

$$\langle \nabla^2 p(z,t),\varphi(t)\rangle - \langle L_\gamma(t) * p(z,t),\varphi(t)\rangle - \left\langle \frac{1}{c_0^2}\frac{\partial^2 p(z,t)}{\partial t^2},\varphi(t)\right\rangle = 0. \tag{12}$$

We are now in a position to investigate the properties of the loss distribution $L_\gamma(t)$ and, consequently, the ultrasonic parameters of our system. Because the loss distribution is related to the dissipation in the system we assume that it is related to the ultrasonic attenuation coefficient. In addition, we assume that this convolution term governs the dispersion of our system as well in order to insure causality. With these assumed properties of the loss distribution we begin by considering the complex ultrasonic wave number. More specifically, we consider the *propagation factor* $\gamma(\omega)=i\kappa(\omega)$ written as

$$\gamma(\omega) = -\alpha(\omega) + ik(\omega), \tag{13}$$

where $\kappa(\omega)$ is the complex wave number, $\alpha(\omega)$ is the attenuation coefficient, $k(\omega)$ is the real wave number, and ω is angular frequency. We have found the propagation factor a convenient quantity to analyze because it provides ready access to the ultrasonic parameters.[5] [Note: In previous publications,[5] we considered the factor $i\kappa(\omega)$, but did not denote it as $\gamma(\omega)$.] Our approach and notation is deliberately similar to that of Szabo[1] in order to facilitate comparisons between the two approaches.

A. The general case

Consider the complex tempered distribution $\gamma(\omega) = -\alpha(\omega)+ik(\omega)$. We relate the real wave number $k(\omega)$ to the attenuation coefficient $\alpha(\omega)$ via the generalized Hilbert transform [Eq. (8)]. Specifically,

$$\langle k(\omega),\varphi(\omega)\rangle=\langle -\mathrm{H}(-\alpha(\omega)),\varphi(\omega)\rangle \tag{14a}$$

$$=\left\langle \frac{(i\omega)^n}{\pi}\left[\frac{-\alpha(\omega)}{(i\omega)^n}*P\left(\frac{1}{\omega}\right)\right],\varphi(\omega)\right\rangle \tag{14b}$$

$$=\int_{-\infty}^{\infty}d\omega\,\varphi(\omega)\frac{(i\omega)^n}{\pi}\int_{-\infty}^{\infty}\frac{\alpha(\omega')}{(i\omega')^n}\frac{d\omega'}{\omega'-\omega} \tag{14c}$$

for some integer n. We note that the generalized Hilbert transform of Eq. (14) is the fundamental Kramers–Krönig dispersion relation relating the attenuation coefficient to the real wave number. We can then rewrite the propagation factor [Eq. (13)] as

$$\gamma(\omega)=-(1-i\mathrm{H})\alpha(\omega), \tag{15}$$

keeping in mind that the propagation factor is a tempered distribution that acts upon a testing function. We make the physical assumption that the time-domain loss distribution of Eq. (12) is given by the inverse Fourier transform of the propagation factor [Eq. (15)],

$$\langle L_\gamma(t),\varphi(t)\rangle\equiv\langle \mathrm{F}^{-1}(\gamma(\omega)),\varphi(t)\rangle. \tag{16}$$

Substituting Eq. (15) into Eq. (16), one finds

$$\langle L_\gamma(t),\varphi(t)\rangle=\langle \mathrm{F}^{-1}[-(1-i\mathrm{H})\alpha(\omega)],\varphi(t)\rangle \tag{17a}$$

$$=\langle -(1+\mathrm{sgn}(t))\mathrm{F}^{-1}(\alpha(\omega)),\varphi(t)\rangle \tag{17b}$$

$$=\langle -2\,1_+(t)\mathrm{F}^{-1}(\alpha(\omega)),\varphi(t)\rangle, \tag{17c}$$

where $1_+(t)$ is the Heaviside step function.[7] We can rewrite Eq. (17) as

$$L_\gamma(t)=-2\,1_+(t)\mathrm{F}^{-1}(\alpha(\omega)) \tag{18}$$

with the implicit assumption of the duality notation. We immediately recognize that the validity of the Hilbert transform for a tempered distribution in the frequency domain insures causality, as has been observed many others.[4,10,14,16] By that we mean any ultrasonic response always follows that of a stimulus. We now have a time-domain loss distribution that is simply related to the inverse Fourier transform of the attenuation coefficient.

B. A special case: Attenuation obeying a frequency power law ($1<y<2$)

Consider those materials (e.g., tissue, plastics, liquids) whose attenuation is governed by Eq. (1) with $1<y<2$ typically. Our goal now is to determine the form of the dispersion as predicted via the generalized Hilbert transform [Eq. (14)], or equivalently the Kramers–Krönig dispersion relations. In addition, we want to determine the time-domain representation of the propagation factor $\gamma(\omega)$. Inserting Eq. (1) into Eq. (14), one can show

$$\langle k(\omega),\varphi(\omega)\rangle=\langle -\mathrm{H}(-\alpha(\omega)),\varphi(\omega)\rangle \tag{19a}$$

$$=\left\langle -\alpha_0\frac{(i\omega)^2}{\pi}\left[\frac{|\omega|^y}{(i\omega)^2}*P\left(\frac{1}{\omega}\right)\right],\varphi(\omega)\right\rangle \tag{19b}$$

$$=\frac{\alpha_0}{\pi}\int_{-\infty}^{\infty}d\omega\,\varphi(\omega)\omega^2\int_{-\infty}^{\infty}\frac{|\omega'|^y}{\omega'^2}\frac{d\omega'}{\omega'-\omega} \tag{19c}$$

$$=\alpha_0\tan\left(\frac{y\pi}{2}\right)\int_{-\infty}^{\infty}d\omega\,\varphi(\omega)\omega^y \tag{19d}$$

$$=\left\langle \alpha_0\tan\left(\frac{y\pi}{2}\right)\omega^y,\varphi(\omega)\right\rangle \tag{19e}$$

where we have used the value $n=2$ for the generalized Hilbert transform. If we carry this through to phase velocity and evaluate at positive frequencies ω_1 and ω_2, we find

$$\left\langle \frac{1}{c(\omega_2)}-\frac{1}{c(\omega_1)},\varphi(\omega)\right\rangle$$

$$=\left\langle \alpha_0\tan\left(\frac{y\pi}{2}\right)(\omega_2^{y-1}-\omega_1^{y-1}),\varphi(\omega)\right\rangle. \tag{20}$$

We recognize the dispersion

$$\frac{1}{c(\omega_2)}-\frac{1}{c(\omega_1)}=\alpha_0\tan\left(\frac{y\pi}{2}\right)(\omega_2^{y-1}-\omega_1^{y-1}), \tag{21}$$

which has been reported previously. [See Eq. (10) of Ref. 5.]

We conclude this section with a look at the time-domain loss distribution $L_\gamma(t)$. As we have seen in Sec. III A the loss distribution is simply related to the inverse Fourier transform of the attenuation coefficient distribution. We find [see p. 43, Table 1 of Ref. 13 or p. 251, Eq. (16) of Ref. 17]

$$\langle L_\gamma(t),\varphi(t)\rangle$$

$$=\langle -21_+(t)\mathrm{F}^{-1}(\alpha(\omega)),\varphi(t)\rangle \tag{22a}$$

$$=\langle -2\alpha_0 1_+(t)\mathrm{F}^{-1}(|\omega|^y),\varphi(t)\rangle \tag{22b}$$

$$=-2\alpha_0\int_{-\infty}^{\infty}dt\,\varphi(t)1_+(t)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}d\omega\,|\omega|^y e^{-i\omega t} \tag{22c}$$

$$=-4\alpha_0 y!\cos\left(\frac{\pi}{2}(y+1)\right)\int_{-\infty}^{\infty}dt\,\varphi(t)\frac{1_+(t)}{|t|^{y+1}} \tag{22d}$$

$$=\left\langle -4\alpha_0 y!\cos\left(\frac{\pi}{2}(y+1)\right)\frac{1_+(t)}{|t|^{y+1}},\varphi(t)\right\rangle. \tag{22e}$$

(We note that the choice of normalization and sign convention for the Fourier transform will affect the numerical constants of the time-domain loss distribution.) We now observe that by assuming the behavior of the attenuation coefficient we can predict the dispersion via the generalized Hilbert transform (or generalized Kramers–Krönig dispersion relations), and we can determine the time-domain loss distribution via the inverse Fourier transform.

IV. REVIEW OF THE TIME-CAUSAL THEORY

The time-causal theory of Szabo begins with a dissipative wave equation similar to Eq. (11). [See Eq. (2) of Ref. 1.] The time-causal theory also assumes an attenuation coefficient given by Eq. (1). Similarly, the time-causal theory assumes the ultrasonic parameters are represented by generalized functions (i.e., tempered distributions) as opposed to

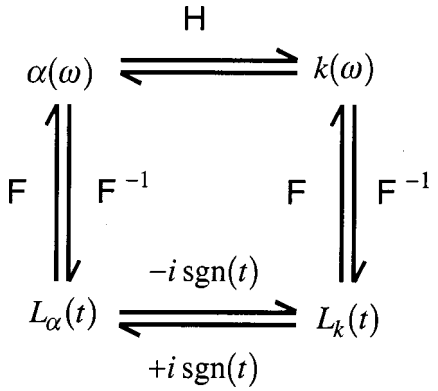


FIG. 1. The relations between the ultrasonic parameters and time-domain loss distributions are shown. The methods to transform between the representations are provided.

point functions. By taking the inverse Fourier transform of the attenuation coefficient Szabo is able to find a time-domain loss distribution. [See Eqs. (37) and (43) of Ref. 1.] In addition, by utilizing the time-domain analog of the frequency-domain Hilbert transform [Eq. (9)], Szabo defines the *time-causal relations* [Eqs. (41) and (42) of Ref. 1] as

$$L_k = -i\,\mathrm{sgn}(t) L_\alpha \tag{23}$$

and

$$L_\alpha = +i\,\mathrm{sgn}(t) L_k\,. \tag{24}$$

Considering the time-domain loss distribution as being composed of Eqs. (23) and (24), Szabo then finds [Eq. (43) of Ref. 1]

$$L_\gamma(t) = L_\alpha(t) + iL_k(t) \tag{25a}$$

$$= (1+\mathrm{sgn}(t))L_\alpha(t) \tag{25b}$$

$$= 21_+(t)L_\alpha(t). \tag{25c}$$

By applying the Fourier transform to Eq. (23), Szabo determines the identical dispersion relation to that of the Kramers–Krönig frequency-domain approach [Eq. (29) of Ref. 18], assuming $1<y<2$.

V. DISCUSSION

We have shown that the time-causal theory and the Kramers–Krönig dispersion relations using tempered distributions are equivalent methods for determining dispersion from an attenuation coefficient of the form of Eq. (1). We have considered the specific case of $1<y<2$, although similar analysis can be performed for other ranges of y.[1,5,18] More significantly, we have shown in Sec. III A that there is no *a priori* restriction to consideration of only attenuation distributions of frequency power laws. One could consider other attenuation distributions assuming valid Fourier and Hilbert transforms (in a distribution sense) are available.

The equivalence of the two approaches is shown in Fig. 1. The individual components of the propagation factor [Eq. (13)] and the time-domain loss distribution [Eq. (25)] are shown. In addition, the methods (Fourier or Hilbert transform) are noted to suggest how to determine a particular quantity given another quantity. Both the time-causal and Kramers–Krönig approaches begin in the upper left-hand corner with the assumed model of the attenuation coefficient. To predict dispersion, the Kramers–Krönig approach follows the clockwise path by applying the Hilbert transform to the attenuation coefficient. Alternatively, the time-causal approach initially determines the time-domain loss distribution via the inverse Fourier transform of the attenuation coefficient, and eventually Fourier transforms back into the frequency domain to determine the dispersion.

VI. CONCLUSION

The present work and the time-causal theory comprise complementary views of causality. Specifically, the time-causal loss distribution [Eq. (25)] and the frequency-domain propagation factor [Eq. (15)] are a Fourier transform pair of each other. Development of the frequency-domain (exact) Kramers–Krönig dispersion relations requires knowledge of either the attenuation or phase velocity at all frequencies. The attraction of a time-domain-based approach is that analysis is naturally bounded by the finite duration of the ultrasonic pressures, as Szabo has noted.[1] The development of the present work provides a framework for developing time-domain loss distributions for a broader class of generalized functions than those of power law forms, namely those generalized functions which possess a Fourier transform. In an effort to demonstrate the unity of these two approaches, we have tried to show how all the pieces can be integrated, building on our previous work[5] and the significant earlier contributions of Szabo.[1,18]

ACKNOWLEDGMENTS

The authors would like to thank Steven L. Baldwin for careful review of the manuscript.

[1] T. L. Szabo, ''Time domain wave equations for lossy media obeying a frequency power law,'' J. Acoust. Soc. Am. **96**, 491–500 (1994).
[2] R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, 1st ed. (American Mathematical Society, New York, 1934), pp. 16–19.
[3] A. Papoulis, *The Fourier Integral and its Applications*, reissued ed. (McGraw–Hill, New York, 1987), pp. 215–217.
[4] H. M. Nussenzveig, *Causality and Dispersion Relations*, 1st ed. (Academic, New York, 1972), pp. 28–43.
[5] K. R. Waters, M. S. Hughes, J. Mobley, G. H. Brandenburger, and J. G. Miller, ''On the applicability of Kramers-Krönig relations for media with ultrasonic attenuation obeying a frequency power law,'' J. Acoust. Soc. Am. **108**, 556–563 (2000).
[6] E. Pfaffelhuber, ''Generalized impulse response and causality,'' IEEE Trans. Circuit Theory **CT-18**(2), 218–223 (1971).
[7] A. H. Zemanian, *Distribution Theory and Transform Analysis: An Introduction to Generalized Functions, with Applications* (Dover, New York, 1987), pp. 2–14, 185.
[8] W. Güttinger, ''Generalized functions and dispersion relations in physics,'' Fortschr. Phys. **14**, 483–602 (1966).
[9] L. Schwartz, *Mathematics for the Physical Sciences*, revised ed. (Addison–Wesley, Reading, MA, 1966), p. 119.
[10] E. J. Beltrami and M. R. Wohlers, *Distributions and the Boundary Values of Analytic Functions* (Academic, New York, 1966), pp. 42–75.
[11] I. M. Gel'fand and G. E. Shilov, *Generalized Functions*, Vol. I (Academic, New York, 1964), pp. 1–18.
[12] J. N. Pandey, *The Hilbert Transform of Schwartz Distributions and Applications* (Wiley, New York, 1996), pp. 114–136.
[13] M. J. Lighthill, *Introduction to Fourier Analysis and Generalised*

Functions, 1st ed. (Cambridge U.P., Cambridge, 1958), pp. 15–43.
[14] G. Arfken, *Mathematical Methods for Physicists*, 3rd ed. (Academic, Boston, 1985), p. 401.
[15] R. N. Bracewell, *The Fourier Transform and Its Applications*, 2nd revised ed. (McGraw–Hill, New York, 1986), p. 267.
[16] E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, 2nd ed. (Clarendon, Oxford, 1948), pp. 119–131.
[17] M. A. Telléz, R. A. Cerutti, and S. E. Trione, ''Tables of Fourier, Laplace and Hankel transforms of *n*-dimensional generalized functions,'' Acta Applic. Math. **48**(3), 235–284 (1997).
[18] T. L. Szabo, ''Causal theories and data for acoustic attenuation obeying a frequency power law,'' J. Acoust. Soc. Am. **97**, 14–24 (1995).

The effect of the physical properties of the tube wall on the attenuation of sound in evaporating and condensing gas–vapor mixtures

William V. Slaton,[a] Richard Raspet, and Craig J. Hickey
National Center for Physical Acoustics, University of Mississippi, University, Mississippi 38677

(Received 27 September 1999; accepted for publication 12 August 2000)

An investigation of sound propagation in an air–water vapor mixture contained in a cylindrical tube with wet walls was recently completed [Hickey *et al.*, J. Acoust. Soc. Am. **107**, 1126–1130 (2000)]. A generalization to include the heat flux at the tube wall is presented here. The attenuation of sound in air–water vapor mixtures can be affected by the thermal properties of the tube wall. The controlling parameter is ε_s, which is a proportionality constant that relates the heat flux per degree Kelvin for the substrate to that of the gas mixture. For a given amount of heat, provided by expansion and rarefaction of the working fluid, different substrates will undergo different temperature excursions. These temperature swings at the boundary change the vapor pressure of the condensate and thus reduce the diffusion of vapor to and from the boundary resulting in a reduction of the attenuation. © *2000 Acoustical Society of America.* [S0001-4966(00)03411-1]

PACS numbers: 43.35.Ud [HEB]

I. INTRODUCTION

Mao[1] suggested that the inclusion of temperature fluctuations in the walls of a wet tube containing a gas–vapor mixture may increase the attenuation. We have extended our recent calculations[2,3] of attenuation in wet tubes to incorporate the appropriate boundary conditions. Also, Hiller and Swift's recent investigation[4] of condensation in an open flow thermoacoustic refrigerator provides further motivation for developing a detailed understanding of the physics of sound propagation in wet pores for possible application to thermoacoustic systems.

The calculation presented in this article includes dissipation of acoustic energy due to diffusion, thermal relaxation, and viscous relaxation. The tube wall is assumed wet so as to allow mass transfer of vapor to and from the wall. We assume the wet layer is thick enough that it is not evaporated away in an acoustic cycle and yet thin enough so that it is much thinner than a thermal penetration depth in the layer. In a previous article, it was assumed that the temperature of the tube wall and the layer of liquid did not fluctuate. In this article we generalize the model to include temperature perturbation of the wall. This requires the introduction of the latent heat of evaporation/condensation as a source of heat at the tube wall. Solution of this system results in a theory of sound attenuation in gas–vapor mixtures in realistic wet tubes. Table I displays the variables and symbols used in the theoretical development.

A. New boundary conditions

Generalization of the model used in Ref. 3 to include temperature fluctuations at the tube wall introduces the following boundary conditions. We demand that the temperature in the tube wall, T_w, and the gas mixture, T, be equal at the tube wall ($\eta=1$). This translates into a statement about the dimensionless fluctuations of these quantities,

$$T^*=T_w^*. \tag{1}$$

Equation (1) replaces $T^*=0$ at the tube wall assumed in Ref. 3. Allowing the temperature to fluctuate in the tube wall requires an additional boundary condition on the heat flux. In general, this can be written as

$$\vec{q}=\kappa\vec{\nabla}T-\kappa_w\vec{\nabla}T_w, \tag{2}$$

where $\vec{q}$ is the heat flux generated at the boundary. For a system undergoing a phase change at the wall the heat flux generated is written $\vec{q}=l\vec{m}_{\text{flux}}=l\rho_2^0\vec{v}_2(1)$, where ρ_2^0 is the ambient vapor density, $\vec{v}_2(1)$ is the radial velocity of the vapor evaluated at the wall, and l is the latent heat of vaporization. Thus, our boundary condition for the heat flux becomes

$$l\rho_2^0\vec{v}_2(1)=\kappa\vec{\nabla}T-\kappa_w\vec{\nabla}T_w. \tag{3}$$

Employing the low reduced frequency approximation to Eq. (3) and making use of the definition of the radial difference velocity of Ref. 3, V, yields the following equation for the heat flux in the radial direction:

$$-l\rho_2^0\frac{cVR}{T_0}\bigg|_{\eta=1}=\left[\kappa\frac{\partial T^*}{\partial\eta}-\kappa_w\frac{\partial T_W^*}{\partial\eta}\right]_{\eta=1}. \tag{4}$$

Temperature fluctuations at the tube wall induce vapor pressure fluctuations. The Clausius–Clapeyron equation,[5]

$$P_{\text{Vapor}}|_{\eta=1}=P_0\exp\left[-\frac{lm_2}{\Re}\left(\frac{1}{T}-\frac{1}{T_{\text{REF}}}\right)\right]_{\eta=1}, \tag{5}$$

relates the vapor pressure to the temperature. In the above expression, P_0 is 1.01×10^5 Pa and $T_{\text{REF}}=373.16$ K. An expression for the density of the vapor near the tube wall can be derived from the ideal gas law and Eq. (5):

[a]Electronic mail: wmslaton@meta3.net

TABLE I. Variables and symbols used in equations.

ρ	density of the mixture
p	pressure of the mixture
T	temperature of the mixture
$\vec{v}_1$, $\vec{v}_2$	hydrodynamic velocity of the gas, of the vapor
V	radial component of $(\vec{v}_1-\vec{v}_2)/c$
m_1, m_2	molecular mass of the gas, of the vapor
n_1, n_2	number density of the gas, of the vapor
$n=n_1+n_2$	total number density
k_T	thermal diffusion ratio
c_p	specific heat at constant pressure of the mixture
l	latent heat of vaporization
$\Re$	universal gas constant
γ	ratio of specific heats of the mixture
μ	viscosity of the mixture
κ	thermal conductivity of the mixture
D_{12}	mass diffusion coefficient
c	speed of sound in the mixture
R	tube radius
ω	circular frequency
$\xi=\frac{\omega z}{c}$	nondimensional axial variable
$\eta=\frac{r}{R}$	nondimensional radial variable
$\Omega=\frac{\omega R}{c}$	reduced frequency
$\lambda_\mu=R\sqrt{\frac{\rho\omega}{\mu}}$	shear wave number
$\lambda_T=R\sqrt{\frac{\rho\omega c_p}{\kappa}}$	thermal wave number
$\lambda_D=R\sqrt{\frac{\omega}{D_{12}}}$	diffusion wave number

$$\rho_2|_{\eta=1}=\rho_2^0\frac{T_{\mathrm{REF}}}{T}\exp\left[-\frac{lm_2}{\Re}\left(\frac{1}{T}-\frac{1}{T_{\mathrm{REF}}}\right)\right]_{\eta=1}. \tag{6}$$

Linearizing with $T=T_0(1+T^*e^{-i\omega t})$ yields

$$\rho_2|_{\eta=1}=\rho_2^0\left(1+\left(\frac{lm_2}{\Re T_0}-1\right)T^*e^{-i\omega t}\right)_{\eta=1}, \tag{7}$$

where T_0 is the ambient temperature of the system. The equation of continuity can be used to write Eq. (7) in terms of ρ^*, V, and T^*:

$$i\Omega\rho^*|_{\eta=1}=\left[\frac{\rho_1^0}{\rho^0}\frac{1}{\eta}\frac{\partial}{\partial\eta}(\eta V)+i\Omega\left(\frac{lm_2}{\Re T_0}-1\right)T^*\right]_{\eta=1}. \tag{8}$$

Use of the equation of state then yields

$$i\Omega P^*|_{\eta=1}=\left[\frac{n_1}{n}\frac{1}{\eta}\frac{\partial}{\partial\eta}(\eta V)+i\Omega\left(\frac{lm_2}{\Re T_0}\right)T^*\right]_{\eta=1}. \tag{9}$$

Equation (9) replaces

$$i\Omega P^*|_{\eta=1}=\left[\frac{n_1}{n}\frac{1}{\eta}\frac{\partial}{\partial\eta}(\eta V)\right]_{\eta=1} \tag{10}$$

as the boundary condition on pressure at the tube wall from Ref. 3.

The remaining boundary conditions of Ref. 3 still apply to this system. These include no slip for the mixture at the tube wall and the hydrodynamic and radial difference velocities must be zero on the tube axis since the axis is not a source or sink. Also, a boundary condition unchanged from Ref. 3 is that at the tube wall the inert gas component cannot enter the wall,

$$v_1=0, \tag{11}$$

which translates to a boundary condition on v and V as

$$v+\frac{\rho_2}{\rho}V=0. \tag{12}$$

Using Eq. (8) with the equation of continuity from Ref. 3 and the equation for the axial hydrodynamic velocity yields an expression that when multiplied through by $\eta\partial/\partial\eta$, integrating from 0 to 1 and simplified using Eq. (12), is

$$\left[\frac{i\Omega}{2\gamma}\frac{\partial^2P^*}{\partial\xi^2}F(\lambda_\mu)+\frac{i\Omega}{2}P^*+\frac{n_2}{n}V\right]_{\eta=1}=i\Omega\int_0^1T^*\eta\,d\eta. \tag{13}$$

Thus, Eqs. (1), (4), (9), and (13) are the new boundary conditions used in this article.

The equation for the temperature distribution in the tube wall is

$$\frac{\partial T_w}{\partial t}-\frac{\kappa_w}{c_{pw}\rho_w}\nabla^2T_w=0, \tag{14}$$

where T_w is the temperature, ρ_w is the density, c_{pw} is the specific heat capacity at constant pressure, and κ_w is the thermal conductivity of the tube wall. The physical properties of several substrates are listed in Table II.[6] The choice of these substrates is motivated by an interest in investigating limiting cases of wall properties. Tube walls constructed of steel remove heat quickly, allowing little temperature fluctuations to affect the vapor pressure, whereas walls constructed of cork hold heat and modify the vapor pressure significantly. However, the porous nature of wood and cork limit their use in a real system. The inclusion of water as a wall material in Table II is appropriate because it represents the limiting case of a condensed layer that is much thicker than a thermal penetration depth in the layer.

TABLE II. Substrate properties at 10 kHz and 373 K.

Substrate	ρ_w (kg/m^3)	κ_w (mW/K m)	c_{pw} (J/g K)	$\lvert\varepsilon_s\rvert$
Steel 304	7900	14000	0.477	0.001 01
Glass	2600	4100	0.81	0.002 50
Water	988	643.5	4.18	0.004 52
Wood (Fir)	420	110	2.7	0.020 88
Cork	160	43	1.9	0.064 50

Following the low reduced frequency method as used in Ref. 2, Eq. (14) assumes the form

$$T_w^*-\frac{i}{\lambda_w^2}\left[\frac{\partial^2T_w^*}{\partial\eta^2}+\frac{1}{\eta}\frac{\partial T_w^*}{\partial\eta}\right]=0, \tag{15}$$

for the dimensionless temperature fluctuations in the tube wall, where

$$\lambda_w = R\sqrt{\frac{\rho_w \omega c_{pw}}{\kappa_w}} \tag{16}$$

is the dimensionless thermal wave number in the tube wall.

B. Solution

The solution for the acoustic mode follows the approach of Ref. 3, except our new boundary conditions are used. First, the solution for the temperature fluctuation in the tube wall is

$$T_w^* = EH_0^{(1)}(\sqrt{i}\lambda_w \eta), \tag{17}$$

where $H_0^{(1)}$ is the Hankel function of the first kind of order zero and E is some constant. This form of the solution is chosen to ensure that the functional form of the temperature fluctuation in the tube wall decays to zero as η becomes large.

The form of the temperature fluctuation in the gas mixture and the radial difference velocity are unchanged from Ref. 3. These are

$$T^* = AJ_0(\sqrt{i}^*\lambda_T \eta) - iD\frac{k_T}{\Omega}\frac{\gamma-1}{\gamma}\frac{\sqrt{i}^*\lambda_D}{1-{}^*\lambda_D^2/\lambda_T^2} \times J_0(\sqrt{i}^*\lambda_D \eta) + \frac{\gamma-1}{\gamma}P^* \tag{18}$$

and

$$V = A\frac{\Omega k_T}{\lambda_D^2}\frac{n^2}{n_1 n_2}\frac{\sqrt{i}^*\lambda_T}{1-{}^*\lambda_T^2/\lambda_D^2}J_1(\sqrt{i}^*\lambda_T \eta) + DJ_1(\sqrt{i}^*\lambda_D \eta), \tag{19}$$

where ${}^*\lambda_T$ and ${}^*\lambda_D$ are solutions to the eigenvector equation,

$$\left(1-\frac{\lambda_x^2}{\lambda_T^2}\right)\left(1-\frac{\lambda_x^2}{\lambda_D^2}\right) = \frac{n^2}{n_1 n_2}k_T^2\frac{\gamma-1}{\gamma}\frac{\lambda_x^2}{\lambda_D^2}. \tag{20}$$

The solutions, ${}^*\lambda_T$ and ${}^*\lambda_D$, result as a consequence of the coupling between the entropy and diffusion equations and can be found in Ref. 3.

The application of Eqs. (17)–(19) to the boundary conditions, Eqs. (1), (4), (9), and (13), yields four equations for the four unknowns A, D, E and $d^2P^*/d\xi^2$ in terms of the dimensionless pressure fluctuation, P^*. Eliminating E between Eqs. (1) and (4) and using the thermoviscous dissipation function,[7]

$$F(\lambda_j) = 1 - \frac{2J_1(\sqrt{i}\lambda_j)}{\sqrt{i}\lambda_j J_0(\sqrt{i}\lambda_j)}, \tag{21}$$

where j denotes T, D, or μ, leads to the following matrix equation:

$$\begin{pmatrix} i\Omega\varphi\frac{\gamma}{\gamma-1}J_0(\sqrt{i}^*\lambda_T)S_1 & \frac{n_1}{n}\sqrt{i}^*\lambda_D J_0(\sqrt{i}^*\lambda_D)S_2 & 0 \\ -J_0(\sqrt{i}^*\lambda_T)S_5 & -\sqrt{i}^*\lambda_D J_0(\sqrt{i}^*\lambda_D)\frac{S_6}{i\Omega} & 0 \\ -i\Omega J_0(\sqrt{i}^*\lambda_T)(1-F({}^*\lambda_T))S_3 & \frac{n_2}{n}\sqrt{i}^*\lambda_D J_0(\sqrt{i}^*\lambda_D)(1-F({}^*\lambda_D))S_4 & \frac{i\Omega}{\gamma}F(\lambda_\mu) \end{pmatrix} \cdot \begin{pmatrix} A \\ D \\ \frac{d^2P^*}{d\xi^2} \end{pmatrix} = \begin{pmatrix} i\Omega(1-\varphi)P^* \\ \frac{\gamma-1}{\gamma}P^* \\ -\frac{i\Omega}{\gamma}P^* \end{pmatrix}, \tag{22}$$

where

$$S_1 = 1 + \frac{n}{n_2}\frac{k_T}{\varphi}\frac{\gamma-1}{\gamma}\frac{{}^*\lambda_T^2/\lambda_D^2}{1-{}^*\lambda_T^2/\lambda_D^2}, \tag{23}$$

$$S_2 = 1 + \varphi\frac{n}{n_1}\frac{k_T}{1-{}^*\lambda_D^2/\lambda_T^2}, \tag{24}$$

$$S_3 = 1 - k_T\frac{n}{n_1}\frac{{}^*\lambda_T^2/\lambda_D^2}{1-{}^*\lambda_T^2/\lambda_D^2}, \tag{25}$$

$$S_4 = 1 - \frac{n}{n_2}\frac{\gamma-1}{\gamma}\frac{k_T}{1-{}^*\lambda_D^2/\lambda_T^2}, \tag{26}$$

$$S_5 = 1 + \varepsilon_s\left\{1 - \frac{n^2}{n_1 n_2}\frac{D_{12}l\rho_2^0}{\kappa_g T_0}\frac{k_T}{1-{}^*\lambda_T^2/\lambda_D^2}\right\}, \tag{27}$$

$$S_6 = \frac{\gamma-1}{\gamma}\frac{k_T}{1-{}^*\lambda_D^2/\lambda_T^2} + \varepsilon_s\frac{{}^*\lambda_D^2}{{}^*\lambda_T^2}\frac{1-F({}^*\lambda_D)}{1-F({}^*\lambda_T)} \times\left\{\frac{\gamma-1}{\gamma}\frac{k_T}{1-{}^*\lambda_D^2/\lambda_T^2} - \frac{D_{12}l\rho_2^0}{\kappa_g T_0}\frac{{}^*\lambda_T^2}{\lambda_T^2}\right\}. \tag{28}$$

Solving the system for d^2P^*/dz^2 and recalling that $\xi = \omega z/c$ gives

$$\frac{d^2P^*}{dz^2} = -k^2P^*, \tag{29}$$

where

$$k^2=\frac{\omega^2}{c^2}\frac{1}{F(\lambda_\mu)}\Bigg\{1+\gamma\frac{n_2}{n}\frac{S_4}{S_7}(1-F(*\lambda_D))[(1-\varphi)S_5+\varphi S_1]$$
$$+\gamma\frac{S_3}{S_7}(1-F(*\lambda_T))\left[(1-\varphi)S_6+\frac{\gamma-1}{\gamma}\frac{n_1}{n}S_2\right]\Bigg\}, \quad (30)$$

with

$$S_7=\frac{n_1}{n}S_2S_5-\varphi\frac{\gamma}{\gamma-1}S_1S_6, \quad (31)$$

$$\varphi=\frac{l}{c_pT_0}, \quad (32)$$

$$\varepsilon_s=\frac{\kappa^*\lambda_T}{\kappa_w\lambda_w}\frac{J_1(\sqrt{i}^*\lambda_T)H_0^{(1)}(\sqrt{i}\lambda_w)}{J_0(\sqrt{i}^*\lambda_T)H_1^{(1)}(\sqrt{i}\lambda_w)}, \quad (33)$$

where φ is the latent heat parameter as defined by Marble[8] and ε_s was introduced by Swift[9] to include temperature fluctuations at the gas–wall interface in his analysis of thermoacoustic engines. For small k_T systems ε_s can be written as

$$\varepsilon_s=-\frac{\sqrt{\rho c_p\kappa\omega}}{\sqrt{\rho_w c_{pw}\kappa_w\omega}}\frac{J_1(\sqrt{i}\lambda_T)H_0^{(1)}(\sqrt{i}\lambda_w)}{J_0(\sqrt{i}\lambda_T)H_1^{(1)}(\sqrt{i}\lambda_w)}. \quad (34)$$

The square root terms in Eq. (34) can be seen to have units of W/m^2 K. Thus, ε_s can be considered as a ratio of the thermal effects of the mixture to those of the tube wall. The larger (smaller) ε_s, the more (less) the temperature of the tube wall will follow the temperature of the gas mixture. The magnitude of ε_s for various substrates has been listed in Table I and can be seen to vary from $O(10^{-3})$ for steel to $O(10^{-2})$ for cork.

Equation (30) is the principle result of this article and is a simple form easily compared to previous results without mass transfer and without the additional boundary conditions on temperature. The usual form of the wave number squared for sound propagation in tubes with viscosity and thermal losses is recovered by setting $n_2=0$, $k_T=0$, and $\varepsilon_s=0$. Limiting forms of the general expression Eq. (30) lead to expressions obtained by Ref. 2 where the limit ε_s and k_T go to zero and Ref. 3 where the limit ε_s goes to zero.

II. AIR–WATER VAPOR MIXTURES

Consider a hypothetical experiment consisting of a rigid tube connected to a source of water vapor and to a source of gas. Starting with the tube under vacuum conditions, allow water vapor to enter the tube. After equilibrium is reached, disconnect from the water vapor source and slightly reduce the temperature of the tube so that some of the vapor condenses on the tube wall. This assures that the assumption of a wet tube wall is satisfied. The amount of water vapor and the water vapor pressure in the tube are functions of the tube temperature. Various amounts of gas can be placed in the tube by connecting to a source of gas. The ideal gas law can be used to determine the number of moles of gas in the tube from measurements of the gas pressure and the tube temperature. The total pressure inside the tube is the sum of the vapor pressure and gas pressure.

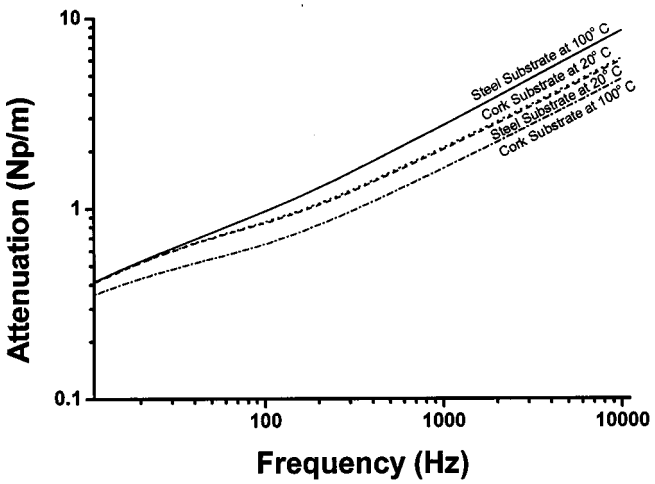


FIG. 1. Predictions of attenuation of sound in a $\frac{1}{2}$-mm-radius tube filled with an air–water vapor mixture for tubes made of steel and cork at 20 °C and 100 °C; steel at 20 °C (dashed line), cork at 20 °C (dotted line), steel at 100 °C (solid line), and cork at 100 °C (dash-dot line). The ambient air pressure is 1 atm.

Calculating the attenuation of sound from the above equations requires values for the properties of the components as well as values for the effective properties of the mixture. The ratio of specific heats and the specific heat at constant pressure for the components are independent of temperature and are calculated from basic kinetic theory, whereas the temperature dependence of the ambient sound speed of the components is taken into account. The ratio of specific heats, the specific heat at constant pressure, and the temperature dependence of ambient sound speed of the mixture are derived based upon an average energy of the mixture. The temperature dependence of the viscosity and thermal conductivity of the components and of the binary mixture are accounted for. The temperature and pressure dependence of mass diffusion and the thermal diffusion ratio of the binary mixture are also accounted for. Reference 2 contains more information concerning the temperature and pressure dependence of these properties.

Example calculations of the attenuation of sound in a tube filled with an air–water vapor mixture using the complete solution, Eq. (30), are shown in Fig. 1. The tube radius is 0.5 mm and the air pressure is set at 1.0 atm (1.01×10^5 Pa). At an ambient temperature of 20 °C there is virtually no difference in the predicted attenuation for a tube constructed of steel or cork. However, at 100 °C ambient temperature the differences in attenuation for the different substrates become more apparent. The larger number of water molecules present in the gas mixture at 100 °C that participate in evaporation and condensation can account for this difference. For this air–water vapor mixture, the frequency range of 10 Hz to 10 kHz represents a range in shear wave number, λ_μ, of about 1 to about 35.

Predictions of attenuation of 10 kHz sound as a function of temperature in a wet tube of radius 0.5 mm filled with an air–water vapor mixture is shown in Fig. 2. If the effects of the latent heat and substrate properties are ignored, then there is an increase in the attenuation if mass transfer is included but a decrease in attenuation if the mixture is treated as an

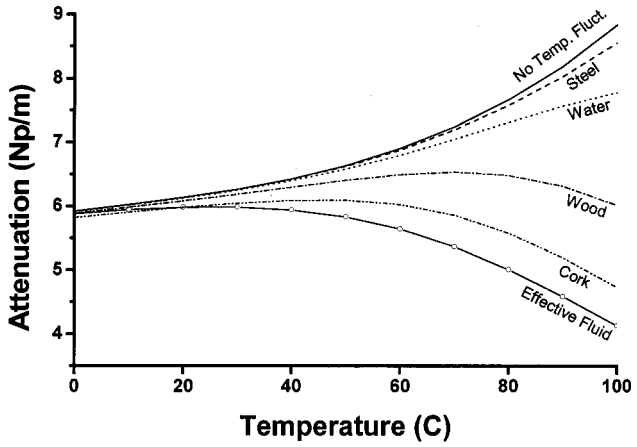


FIG. 2. Predictions of attenuation of 10 kHz sound in a $\frac{1}{2}$-mm-radius tube filled with an air–water vapor mixture for tubes made of various materials. The thermal wave number is of order 10 for this system. The ambient air pressure is 1 atm. No temperature fluctuation (solid line), steel (dashed line), water (dotted line), wood (dash-dot line), cork (dash-dot-dot line), and effective fluid (circle-solid line).

effective fluid, i.e., without mass transfer. When latent heat effects and substrate properties are included it is found that the attenuation decreases with increasing ε_s until it approaches that of the effective fluid. This can be understood from the following description of the physical process involved.

At the tube wall a thin water layer exists to provide (absorb) water vapor during the process of evaporation (condensation). In the analysis of Ref. 3 it was assumed that there were no temperature fluctuations at the tube wall. No temperature fluctuations at this boundary implies that the vapor pressure of the water vapor remains constant since the vapor pressure is a function of temperature only and temperature remains constant at the tube wall. Thus, compression of the gas above the boundary increases the partial pressure of the vapor above the vapor pressure at ambient temperature. This leads to the diffusion of water vapor to the tube wall (condensation). Likewise, on rarefaction of the gas above the boundary the partial pressure of the water vapor is decreased below the vapor pressure at ambient temperature. This will lead to the diffusion of water vapor from the tube wall into the gas mixture (evaporation).

When the effects of nonzero temperature fluctuations at this boundary are taken into account this analysis is modified in the following manner. Nonzero temperature fluctuations at the boundary imply that the vapor pressure of the water vapor will increase (decrease) as the temperature increases (decreases). Thus, compression of the gas above the boundary again increases the partial pressure of the vapor, however, the corresponding temperature increase of the wall raises the vapor pressure at the tube wall. The partial pressure of the vapor in the gas mixture will drive less water vapor to the boundary due to the higher vapor pressure at the wall. Likewise, on rarefaction of the gas above the boundary the partial pressure of the water vapor is lowered; however, the corresponding drop in temperature of the wall lowers the vapor pressure at the boundary. Again, the partial pressure of the vapor in the gas mixture will drive less vapor off the boundary due to the lowered vapor pressure at the wall.

Calculations of attenuation with the thermal diffusion ratio, k_T, set equal to zero demonstrated that the coupling between diffusion and heat flux has little effect on attenuation in the case considered herein. The more realistic boundary condition used in this article has not affected this finding of Ref. 3. Also, the effect of temperature fluctuations at the tube wall only significantly affects the diffusion part of the attenuation for the case considered here. That is, for an effective fluid, where there are only viscous and thermal losses, there is no noticeable change to the attenuation when temperature fluctuations at the tube wall are taken into account for the system described in this article.

III. CONCLUSIONS

The attenuation of sound in wet walled tubes was found to be reduced by the inclusion of nonzero temperature fluctuations at the tube wall. The important parameter determining the size of this effect is ε_s. If this parameter is small, the contribution of the wall to the attenuation will be negligible. Thus, it is adequate to ignore contributions to the attenuation from nonzero temperature fluctuations and thermal diffusion for sound propagation in wet pores at atmospheric pressure in materials where ε_s is $O(10^{-3})$ or smaller. Further work will develop theoretical predictions for the efficiency and power density of thermoacoustic engines with wet pores.

ACKNOWLEDGMENTS

This work is jointly supported by the Office of Naval Research and the USDA-ARS National Sedimentation Laboratory. The authors thank H. E. Bass for general discussions and for sharing his insight on the behavior of gas mixtures, and J. M. Sabatier for support and for introducing us to this interesting topic.

[1] Y. Mao, ''Sound attenuation in a cylindrical tube due to evaporation-condensation,'' J. Acoust. Soc. Am. **104**, 664–670 (1998).
[2] R. Raspet, C. J. Hickey, and J. M. Sabatier, ''The effect of evaporation-condensation on sound propagation in cylindrical tubes using the low reduced frequency approximation,'' J. Acoust. Soc. Am. **105**, 65–73 (1999).
[3] C. J. Hickey, R. Raspet, and W. Slaton, ''Effects of thermal diffusion on the attenuation of sound in evaporating and condensing gas-vapor mixtures in tubes,'' J. Acoust. Soc. Am. **107**, 1126–1130 (2000).
[4] R. A. Hiller and G. W. Swift, ''Condensation in a steady-flow thermoacoustics refrigerator,'' J. Acoust. Soc. Am. (to be published).
[5] F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York, 1965).
[6] F. White, *Heat and Mass Transfer* (Addison-Wesley, Reading, MA, 1991).
[7] W. P. Arnott, H. E. Bass, and R. Raspet, ''General formulation of thermoacoustics for stacks having arbitrarily shaped pore cross sections,'' J. Acoust. Soc. Am. **90**, 3228–3237 (1991).
[8] F. E. Marble, ''Some gasdynamic problems in the flow of condensing vapors,'' Astronautica Acta. **14**, 585–614 (1969).
[9] G. W. Swift, ''Thermoacoustic engines,'' J. Acoust. Soc. Am. **84**, 1145–1180 (1988).

Impedance and admittance matrices of symmetric piezoelectric annular bimorphs and their applications

Sung K. Ha[a] and Young H. Kim
Smart Structures and Materials Laboratory, Department of Mechanical Engineering, Hanyang University, 1271 Sa1-Dong, Ansan, Kyounggi-do, Korea 425-791

(Received 19 May 2000; revised 31 July 2000; accepted 2 August 2000)

The 5×5 impedance and admittance matrices of a symmetric triple-layer piezoelectric annular bimorph (PAB) are presented and the PAB with the piezoelectric layers partially covering the shim layer in harmonic motion is analyzed using the matrices. The electromechanical flow vectors are defined as the mechanical displacements and rotations at the inner and outer boundaries and the electric current of the piezoelectric layer; the conjugate parameters, i.e., the electromechanical effort vectors, are accordingly defined. The impedance matrix which relates the flow vector with the effort vector is divided into three matrices, the mechanical, electrical, and electromechanical coupling impedance matrices; each matrix is represented by a block network with five ports. The resonance and the antiresonance frequencies and the effective electromechanical coupling factors of the PAB excited by the partially covering piezoelectric layers are then calculated for various boundary conditions imposed at the inner and outer surfaces. The numerical results by the impedance matrix are compared with those by the finite element methods; they are in excellent agreement with each other. It is also shown that the matrix representations can be easily applied to the piezoelectric circular bimorph (PCB) with the extended shim layer.
[S0001-4966(00)01711-2]

PACS numbers: 43.38.Ar, 43.38.Fx, 43.40.At, 43.40.Dx [SLE]

I. INTRODUCTION

The piezoelectric bimorphs have been widely used because of their simple structure and high sensitivity of displacement with respect to the exciting electric voltage. In the fields of VCR head tracking actuators, relays, servo valves,[1,2] swing CCD mechanisms,[3] and scanning tunneling microscope tips,[4] the beam types of bimorphs are mostly used. The circular or annular types of bimorphs are used particularly often in other fields, such as buzzers, telephone receivers, loud speakers,[5,6] and droplet ejectors.[7–9] The bimorph consists of two piezoelectric layers, which are polarized normal to the interface. The material properties and geometric dimensions of the two piezoelectric layers are commonly identical to each other, with the opposite polarization as shown in Fig. 1. The symmetric bimorph causes bending deformation under an electric field because one piezoelectric layer produces extension, the other contraction. Often in the practical application, a piezoelectric layer is bonded on each side of a nonpiezoelectric shim layer which is used for enhancing strength and stiffness.

Piezoelectric transducers have resonance and antiresonance modes from an electrical point of view. At the resonance state, the electrical impedance reaches a minimum while the mechanical vibration amplitude becomes very large. On the contrary, at the antiresonance state, the electrical impedance reaches a maximum and low current flows through the transducer. Especially in the application of high-power devices such as ultrasonic motors, it is important for the device to be operated at the antiresonance state rather than at the resonance state because the antiresonance state can provide the same mechanical vibration level with less heat generation and with higher Q-factor than the resonance state.[1] In this respect, the accurate calculation of both the resonance frequencies (RF) and the antiresonance frequencies (AF) is essential in the design of the piezoelectric transducers.

Numerous investigations have been performed on the piezoelectric bimorph. Smits *et al.*[10–12] derived the constituent equations of the piezoelectric bimorphs and calculated the RF and the AF of the cantilevered bimorphs by using the dynamic admittance matrix. Tanaka *et al.*[13–15] formulated basic equations of piezoelectric beams from Hamilton's principle and presented a block equivalent circuit of the multilayered piezoelectric benders. Wang *et al.*[16] discussed the electromechanical coupling factors and the output efficiencies of the cantilevered bimorph and unimorph actuators under the static conditions. Cho *et al.*[17] presented a five-port equivalent electric circuit of the triple-layer piezoelectric bimorphs, from which the resonance and the antiresonance frequencies can be determined.

Relatively less investigation has been performed on the annular types of the triple layer bimorph. Perçin *et al.*[7–9] designed annular types of the droplet ejectors with the piezoelectric ring layer partially covering the shim layer. They used the finite element methods (FEMs) to determine the resonance frequencies of the structures, the input impedance, and the normal displacement of the surface. Dobrucki *et al.*[18] presented FEMs for analyzing an axisymmetric bimorph shell. Hui *et al.*[19] studied the electromechanical static response of the circular multilayered piezoelectric plates. Lee *et al.*[20,21] derived two-dimensional first-order governing

[a]Electronic mail: sungha@email.hanyang.ac.kr

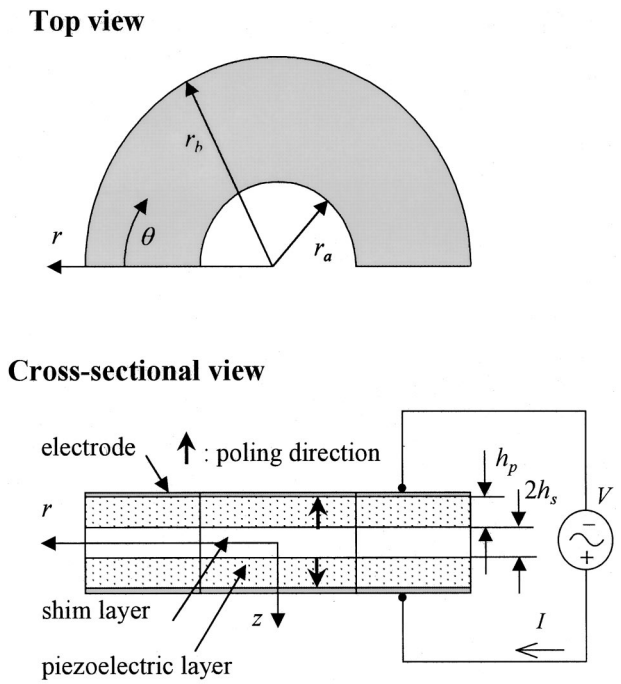


FIG. 1. Configuration of the symmetric triple-layer piezoelectric annular bimorph (PAB) in series connection.

equations for the piezoelectric crystal plates with the thickness-graded material properties and obtained the RF of flexural and thickness-shear vibration. However, the impedance and admittance matrices of a piezoelectric annular bimorph (PAB) and a piezoelectric circular bimorph (PCB) have not been reported yet.

In this article, the equation of motion of the symmetric triple-layer PAB in harmonic motion is derived together with the charge equation. The impedance and admittance matrices of the PAB are then obtained with five ports: two displacements, two rotations, and one electric current. The impedance matrices are also represented by a block network. The RF, the AF, and the effective electromechanical coupling factors (EECF) of the PAB with the piezoelectric layers partially covering the shim layer are calculated by combining the derived matrices, and compared with those calculations by the FEM. The methods are equally applied to the PCB and the effects of the radial size of the piezoelectric layers partially covering the shim layer on the RF and the AF are also studied.

II. EQUATION OF MOTION OF THE PAB

A symmetric triple-layer piezoelectric annular bimorph in series connection is considered in this study, and the top and cross-sectional views in the cylindrical coordinate system r, θ, and z is shown in Fig. 1. The inner radius, the outer radius, the thickness of each piezoelectric layer, and the middle shim layer are denoted by r_a, r_b, h_p, and $2h_s$, respectively. The annular bimorph is assumed to be thin and the top and bottom surfaces are supposed to be stress-free. The z-direction-dependent stresses T_{zz}, T_{zr}, and $T_{z\theta}$ thus vanish throughout. Each piezoelectric layer is polarized in the direction perpendicular to the plate, i.e., parallel to the z axis, and electroded on the top and bottom faces so that the r and θ directional electric fields E_r and E_θ also vanish. All the mechanical and electrical components are assumed to be independent of θ. The two-dimensional constitutive equations for the piezoelectric element are written as[1,22]

$$\begin{aligned} S_{rr} &= s_{11}^E T_{rr} + s_{12}^E T_{\theta\theta} + d_{31} E_z, \\ S_{\theta\theta} &= s_{12}^E T_{rr} + s_{11}^E T_{\theta\theta} + d_{31} E_z, \\ D_z &= d_{31}(T_{rr} + T_{\theta\theta}) + \varepsilon_{33}^T E_z, \end{aligned} \tag{1}$$

which is expressed in a form that is more convenient for the present analysis:

$$\begin{aligned} T_{rr} &= c_{11} S_{rr} + c_{12} S_{\theta\theta} - e_{31} E_z, \\ T_{\theta\theta} &= c_{12} S_{rr} + c_{11} S_{\theta\theta} - e_{31} E_z, \\ D_z &= e_{31}(S_{rr} + S_{\theta\theta}) + \varepsilon_{33} E_z, \end{aligned} \tag{2}$$

where the following relationships hold:

$$\begin{aligned} &c_{11} = \frac{1}{s_{11}^E(1-\nu^2)}, \quad c_{12} = \nu c_{11}, \quad e_{31} = \frac{d_{31}}{s_{11}^E(1-\nu)}, \\ &\varepsilon_{33} = \varepsilon_{33}^T - \frac{2d_{31}^2}{s_{11}^E(1-\nu)}, \quad \nu = -\frac{s_{12}^E}{s_{11}^E}. \end{aligned} \tag{3}$$

In Eqs. (1)–(3), T_{rr}, $T_{\theta\theta}$, S_{rr}, and $S_{\theta\theta}$ are the radial and circumferential normal stresses and strains; D_z and E_z are the electric displacement and field, respectively; s_{11}^E is the elastic compliance under the constant electric field; d_{31} is the piezoelectric strain constant; ε_{33}^T is the dielectric constant under the constant stress: and ν denotes a planar Poisson's ratio.

The mechanical and electrical responses as well as the applied voltages and loads are supposed to be in harmonic motion with a common angular frequency ω. Therefore, the time dependency of all the variables, which can be denoted by $e^{j\omega t}$, is omitted in the following analysis. Based on the Kirchhoff assumption and axisymmetry, the displacements are supposed to be

$$\begin{aligned} &u_r(r,z) = -z\phi_r(r), \quad u_\theta(r,z) = 0, \\ &\phi_r(r) = \frac{\partial u_z}{\partial r}, \quad u_z(r,z) = u_z(r), \end{aligned} \tag{4}$$

where $\phi_r(r)$ denotes the angle of rotation. The strain–displacement relationship with the axial symmetric condition can then be expressed as

$$S_{rr} = \frac{\partial u_r}{\partial r} = -z\frac{\partial \phi_r}{\partial r} = -z\frac{\partial^2 u_z}{\partial r^2}, \tag{5a}$$

$$S_{\theta\theta} = \frac{u_r}{r} = -z\frac{\phi_r}{r} = -\frac{z}{r}\frac{\partial u_z}{\partial r}. \tag{5b}$$

The electric field in the piezoelectric layers is approximated as

$$E_z = \frac{V}{2h_p}, \tag{6}$$

where V is the total electric potential difference and $2h_p$ is the total thickness of the piezoelectric layers.

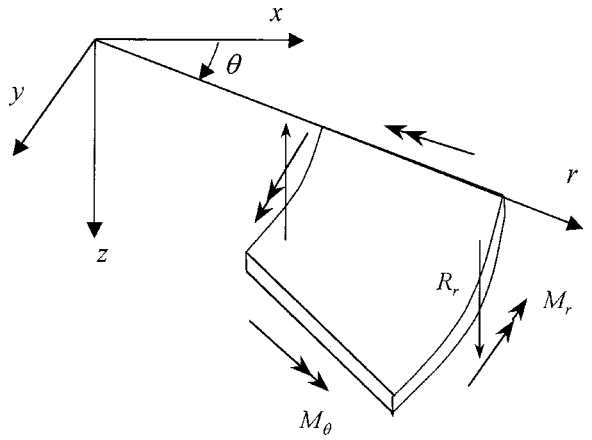


FIG. 2. The flexural moments M_r and M_θ and the shear force R_r acting in the positive directions on the edges.

The flexural moments per unit length, M_r and M_θ, are defined as

$$M_r = -\int T_{rr} z\, dz, \tag{7a}$$

$$M_\theta = -\int T_{\theta\theta} z\, dz. \tag{7b}$$

Substituting Eqs. (2), (4), and (5) into Eq. (7), the flexural moments are expressed as

$$\begin{pmatrix} M_r \\ M_\theta \end{pmatrix} = \begin{bmatrix} D_{11} & D_{12} \\ D_{12} & D_{11} \end{bmatrix} \begin{pmatrix} \partial^2 u_z/\partial r^2 \\ \partial u_z/r\partial r \end{pmatrix} + z_c e_{31} \begin{pmatrix} 1 \\ 1 \end{pmatrix} V, \tag{8}$$

where the flexural stiffness matrix D_{ij} is defined as

$$D_{ij} = \int c_{ij} z^2\, dz = c_{ij}^{(p)}(2h_p z_c^2 + h_p^3/6) + 2c_{ij}^{(s)} h_s^3/3 \tag{9}$$

and the distance z_c denotes the distance from the neutral axis to the middle of the piezoelectric layer, i.e.,

$$z_c = h_s + \frac{h_p}{2}. \tag{10}$$

In Eqs. (9) and (10), the symbols p and s denote the piezoelectric and shim layer, respectively. To satisfy the moment equilibrium, the transverse shear force R_r as shown in Fig. 2 should be given by[23]

$$R_r = -\frac{\partial M_r}{\partial r} - \frac{M_r - M_\theta}{r}$$

$$= -D_{11}\left(\frac{\partial^3 u_z}{\partial r^3} + \frac{1}{r}\frac{\partial^2 u_z}{\partial r^2} - \frac{1}{r^2}\frac{\partial u_z}{\partial r}\right). \tag{11}$$

The differential flexural motion equation can be written as[23]

$$\frac{\partial (rR_r)}{\partial r} = -\rho_0 \omega^2 r u_z. \tag{12}$$

Substitution of Eq. (11) into Eq. (12) yields the motion equation in terms of the displacement:

$$\frac{\partial^4 u_z}{\partial r^4} + \frac{2}{r}\frac{\partial^3 u_z}{\partial r^3} - \frac{1}{r^2}\frac{\partial^2 u_z}{\partial r^2} + \frac{1}{r^3}\frac{\partial u_z}{\partial r} = \lambda^4 u_z, \tag{13}$$

where the parameter λ is defined as

$$\lambda^4 = \frac{\rho_0 \omega^2}{D_{11}} \tag{14}$$

and ρ_0 denotes the mass density per unit area of the bimorph:

$$\rho_0 = \int_{-h_p - h_z}^{h_p + h_z} \rho\, dz = 2(\rho_p h_p + \rho_s h_s), \tag{15}$$

where ρ_p and ρ_s denote the density of the piezoelectric layers and the shim layer, respectively. The general solution of Eq. (13) is[23]

$$u_z(r) = A_1 J_0(\lambda r) + A_2 Y_0(\lambda r) + A_3 I_0(\lambda r) + A_4 K_0(\lambda r), \tag{16}$$

and J_0 and Y_0 are the Bessel functions of the first and second kind of order zero, respectively. Here I_0 and K_0 are modified Bessel functions of the first and second kind of order zero, respectively. The coefficients $A_i (i = 1,2,3,4)$ are to be determined from the boundary conditions.

The electric charge in the piezoelectric layers is defined as

$$Q = \frac{\pi}{h_p}\int \int_{r_a}^{r_b} r D_z\, dr\, dz, \tag{17}$$

which is a conjugate of the voltage when the electric field is defined as in (6). Using Eqs. (2), (5), and (6), the integrand in Eq. (17) can be written as

$$rD_z = -z e_{31}\frac{\partial (r\phi_r)}{\partial r} + r\varepsilon_{33}\frac{V}{2h_p}. \tag{18}$$

Substitution of Eq. (18) into Eq. (17) yields the electric charge Q:

$$Q = 2\pi z_c e_{31}[r_a \phi_r(r_a) - r_b \phi_r(r_b)] + CV, \tag{19}$$

where C denotes the electric capacitance, defined as

$$C = \frac{\pi(r_b^2 - r_a^2)\varepsilon_{33}}{2h_p}. \tag{20}$$

As can be seen in Eq. (19), the electric charge is related with both the rotations at the boundary and the voltage; the electric charge is independent of the translation displacements at the boundary.

III. IMPEDANCE AND ADMITTANCE MATRICES OF PAB AND PCB

In the previous section, the basic equations for the PAB have been derived. In this section, the impedance and admittance matrices will be determined by introducing the lumped parameters. The mechanical effort vector is now defined as

$$\mathbf{F} = (F_1 \quad F_2 \quad F_3 \quad F_4)^T, \tag{21}$$

where F_i $(i = 1,2,3,4)$ denote the shear forces and the flexural moments at $r = r_a$ and r_b, respectively,

$$F_1 = -2\pi r_a M_r|_{r=r_a}, \quad F_2 = -2\pi r_a R_r|_{r=r_a},$$
$$F_3 = 2\pi r_b M_r|_{r=r_b}, \quad F_4 = 2\pi r_b R_r|_{r=r_b}. \tag{22}$$

The mechanical flow vector is also defined as

$$\mathbf{U}=(U_1 \quad U_2 \quad U_3 \quad U_4)^T, \tag{23}$$

where

$$U_1=\left.\frac{\partial \phi_r}{\partial t}\right|_{r=r_a}, \quad U_2=\left.\frac{\partial u_z}{\partial t}\right|_{r=r_a},$$
$$U_3=\left.\frac{\partial \phi_r}{\partial t}\right|_{r=r_b}, \quad U_4=\left.\frac{\partial u_z}{\partial t}\right|_{r=r_b}. \tag{24}$$

The sign conventions of the mechanical flows and efforts are shown in Fig. 3.

Substituting Eq. (16) into Eqs. (8), (11), and (19) yields the mechanical effort and flow vectors and the electric current $I(=j\omega Q)$ can be written as

$$\mathbf{F}=\mathbf{B}_F\mathbf{A}+\mathbf{C}_V V, \tag{25a}$$

$$\mathbf{U}=j\omega\mathbf{B}_U\mathbf{A} \tag{25b}$$

$$I=-\mathbf{C}_V^T\mathbf{U}+j\omega C V, \tag{25c}$$

where the coefficient vector is defined as $\mathbf{A}=(A_1\ A_2\ A_3\ A_4)^T$, and

$$\mathbf{B}_F=2\pi D_{11}\lambda\begin{bmatrix} \lambda r_a J_{0a}-\beta J_{1a} & \lambda r_a Y_{0a}-\beta Y_{1a} & -\lambda r_a I_{0a}+\beta I_{1a} & -\lambda r_a K_{0a}-\beta K_{1a} \\ \lambda^2 r_a J_{1a} & \lambda^2 r_a Y_{1a} & \lambda^2 r_a I_{1a} & -\lambda^2 r_a K_{1a} \\ -\lambda r_b J_{0b}+\beta J_{1b} & -\lambda r_b Y_{0b}+\beta Y_{1b} & \lambda r_b I_{0b}-\beta I_{1b} & \lambda r_b K_{0b}+\beta K_{1b} \\ -\lambda^2 r_b J_{1b} & -\lambda^2 r_b Y_{1b} & -\lambda^2 r_b I_{1b} & \lambda^2 r_b K_{1b} \end{bmatrix}, \tag{26a}$$

$$\mathbf{B}_U=\begin{bmatrix} -\lambda J_{1a} & -\lambda Y_{1a} & \lambda I_{1a} & -\lambda K_{1a} \\ J_{0a} & Y_{0a} & I_{0a} & K_{0a} \\ -\lambda J_{1b} & -\lambda Y_{1b} & \lambda I_{1b} & -\lambda K_{1b} \\ J_{0b} & Y_{0b} & I_{0b} & K_{0b} \end{bmatrix}, \tag{26b}$$

$$\mathbf{C}_V=2\pi e_{31} z_c(-r_a \quad 0 \quad r_b \quad 0)^T, \tag{26c}$$

where J_{1a} denotes $J_1(\lambda r_a)$, etc., and β is defined as $1-D_{12}/D_{11}$.

Rearranging Eq. (25), after eliminating the coefficient vector **A** in Eqs. (25a) and (25b), yields the 5×5 impedance matrix **Z**:

$$\begin{pmatrix} \mathbf{F} \\ V \end{pmatrix}=\mathbf{Z}\begin{pmatrix} \mathbf{U} \\ I \end{pmatrix}, \tag{27}$$

where

$$\mathbf{Z}=\frac{1}{j\omega}\begin{bmatrix} \mathbf{B}_F\mathbf{B}_U^{-1}+\mathbf{C}_V\mathbf{C}_V^T/C & \mathbf{C}_V/C \\ \mathbf{C}_V^T/C & 1/C \end{bmatrix}. \tag{28}$$

The impedance matrix **Z** is now separated into three matrices

$$\mathbf{Z}=\mathbf{Z}_{\mathrm{M}}+\mathbf{Z}_{\mathrm{C}}+\mathbf{Z}_{\mathrm{E}}, \tag{29}$$

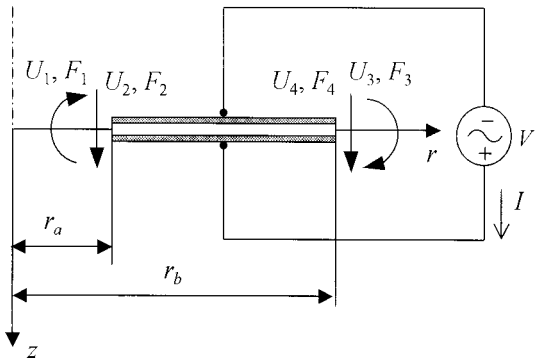


FIG. 3. Flow vector (U_i,I) and effort vector (F_i,V) defined for derivation of the impedance matrix of the PAB.

where $\mathbf{Z}_{\mathrm{M}}$, $\mathbf{Z}_{\mathrm{C}}$, and $\mathbf{Z}_{\mathrm{E}}$ denote the mechanical impedance, electromechanical coupling impedance, and electrical impedance matrices, respectively:

$$\mathbf{Z}_{\mathrm{M}}=\frac{1}{j\omega}\begin{bmatrix} \mathbf{B}_F\mathbf{B}_U^{-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix}, \tag{30a}$$

$$\mathbf{Z}_{\mathrm{C}}=\frac{1}{j\omega C}\begin{bmatrix} \mathbf{C}_V\mathbf{C}_V^T & \mathbf{C}_V \\ \mathbf{C}_V^T & 0 \end{bmatrix}, \tag{30b}$$

$$\mathbf{Z}_{\mathrm{E}}=\frac{1}{j\omega C}\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}. \tag{30c}$$

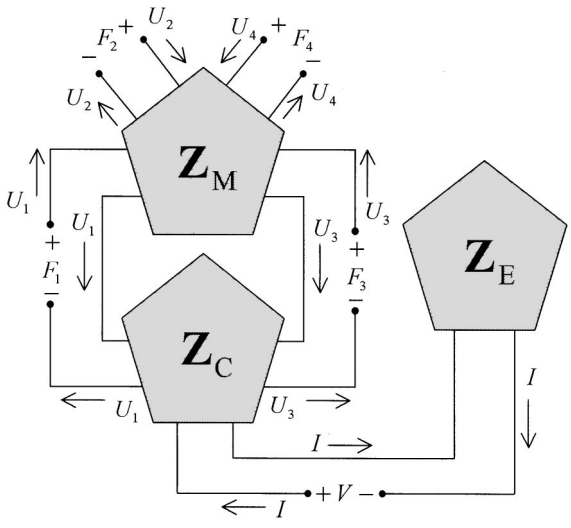


FIG. 4. Five-port impedance matrix of the PAB represented as three block networks in the series connection; a mechanical impedance, an electrical impedance, and an electromechanical coupling impedance.

TABLE I. The comparison of the RF (ω_r) and the AF (ω_a) of the symmetric triple-layer PAB calculated by the admittance matrix and the FEM.

Boundary condition	Mode no.	Equivalent circuit ω_r (Hz)	Equivalent circuit ω_a (Hz)	FEM ω_r (Hz)	FEM ω_a (Hz)	Error (%) ω_r	Error (%) ω_a
Free–Free	1	1590	1706	1587	1721	0.2	0.9
	2	10 326	11 031	10 150	11 020	1.7	0.1
	3	27 032	27 107	25 980	26 490	4.1	2.3
Clamped–Free	1	1386	1517	1396	1548	0.7	2.0
	2	8917	9284	8843	9259	0.8	0.3
	3	25 786	26 140	25 040	25 700	3.0	1.7
Free–Clamped	1	2270	2305	2279	2326	0.4	0.9
	2	10 595	10 684	10 490	10 590	1.0	0.9
	3	27 419	27 521	26 580	26 920	3.2	2.2

The impedance matrix $\mathbf{Z}$ is now represented by a series connection of $\mathbf{Z}_{\mathrm{M}}$, $\mathbf{Z}_{\mathrm{C}}$, and $\mathbf{Z}_{\mathrm{E}}$, as shown in Fig. 4. Notice that they are all symmetric. Each impedance matrix is represented as a polygon that has the five sides where the bottom side corresponds to the electrical port and the rest of four sides correspond to the mechanical ports, i.e., two displacements and two rotations. The electromechanical coupling matrix $\mathbf{Z}_{\mathrm{C}}$ has three active ports, i.e., two ports for rotations and one port for the electrical port since $\mathbf{C}_V$ in Eq. (30b) has the nonzero values at the locations corresponding to the rotation ports, as can be seen in Eq. (26c). The networks illustrated in Fig. 4 are very useful in the analysis of an electromechanical system since the mechanical or electrical ports can be appropriately connected to other ports according to the continuation conditions at the interfaces; the conventional network theory can be effectively used in determining the electromechanical behavior of the electromechanical system.

The admittance matrix can be also obtained by eliminating the coefficient vector $\mathbf{A}$ in Eqs. (25a) and (25b):

$$\begin{pmatrix} \mathbf{U} \\ I \end{pmatrix} = j\omega \begin{bmatrix} \mathbf{B}_U \mathbf{B}_F^{-1} & -\mathbf{B}_U \mathbf{B}_F^{-1} \mathbf{C}_V \\ \text{sym.} & \mathbf{C}_V^T \mathbf{B}_U \mathbf{B}_F^{-1} \mathbf{C}_V + C \end{bmatrix} \begin{pmatrix} \mathbf{F} \\ V \end{pmatrix}, \tag{31}$$

from which the electrical admittance Y can be defined as

$$Y = \frac{1}{V} = j\omega(\mathbf{C}_V^T \mathbf{B}_U \mathbf{B}_F^{-1} \mathbf{C}_V + C). \tag{32}$$

The RF(ω_r) and the AF(ω_a) are then determined by the poles and the zeros of Eq. (32). Using the RF and the AF, the EECF k_{eff} is calculated by applying the known expression[24]

$$k_{\mathrm{eff}}^2 = \frac{\omega_a^2 - \omega_r^2}{\omega_a^2}, \tag{33}$$

which is commonly used as an index of the capability of the piezoelectric transducer to convert mechanical energy into electrical energy, or vice versa, at each RF.

The impedance matrix of a PCB of the radius r_b can be derived in a similar way. In the case of the PCB, the general solution is expressed in terms of only J_0 and I_0 since the displacement has a finite value at $r=0$; the Bessel functions Y_0 and K_0 go to infinity at $r=0$:

$$u_z(x) = A_1 J_0(\lambda r) + A_2 I_0(\lambda r). \tag{34}$$

Following the same procedure as the PAB, $\mathbf{B}_F$ and $\mathbf{B}_U$ in Eqs. (26a) and (26b) now reduce to 2×2 matrices and $\mathbf{C}_V$ in Eq. (26c) reduces to a 2×1 matrix:

$$\mathbf{B}_F = 2\pi D_{11}\lambda \begin{bmatrix} -\lambda r_b J_{0b} + \beta J_{1b} & \lambda r_b I_{0b} - \beta I_{1b} \\ -\lambda^2 r_b J_{1b} & -\lambda^2 r_b I_{1b} \end{bmatrix}, \tag{35a}$$

$$\mathbf{B}_U = \begin{bmatrix} -\lambda J_{1b} & \lambda I_{1b} \\ J_{0b} & I_{0b} \end{bmatrix}, \tag{35b}$$

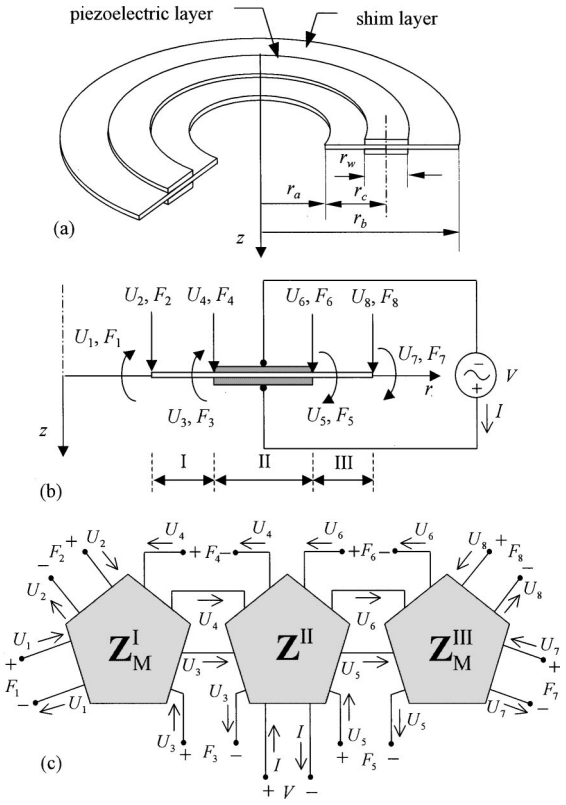


FIG. 5. (a) PAB with the piezoelectric layers partially covering the shim layer; (b) three sections with flow (U_i, I) and effort (F_i, V) vectors; and (c) each section represented by the impedance matrix, all in series connection.

$$\mathbf{C}_V = 2\pi e_{31} z_c (r_b \quad 0)^T, \tag{35c}$$

$$C = \frac{\pi r_b^2 \varepsilon_{33}}{2h_p}, \tag{35d}$$

which yield the 3×3 impedance matrix of Eq. (28) and the admittance matrix of Eq. (31) which have one electrical port and two mechanical ports, i.e., one displacement and one rotation at $r=r_b$.

IV. APPLICATIONS AND DISCUSSION

Several types of PAB are investigated using the impedance and admittance matrices derived in the previous sections. The PAB under consideration consists of PSI-5A-S2 piezoelectric ceramics and a stainless steel layer for the middle shim; the thickness of each layer is $h_p=0.5$ mm and $h_s=0.25$ mm, respectively. The material properties are as follows: for the PSI-5A-S2, $1/s_{11}^E=61$ GPa, $\nu=0.3$, $\rho=7750$ kg/m^3, $d_{31}=0.171$ nm/V, and $\varepsilon_{33}^T=15.3$ nF/m; for the stainless steel, $1/s_{11}^E=200$ GPa, $\nu=0.3$, and $\rho=7830$ kg/m^3.

In the first case, in order to verify the derived impedance and admittance matrices of the PAB, the RF and the AF of the PAB with the radial dimensions of $r_a=10$ mm and $r_b=30$ mm are calculated using Eq. (32) and compared with the results by the three-dimensional FEM. The element has four degrees of freedom at a node, i.e., three displacements and one electric potential;[25] 40 elements are used for the radial direction, and 9 elements for the thickness directions. It is confirmed that the number of elements assures the converged results. Three different boundary conditions are then taken into account, (1) free–free, (2) clamped–free, and (3) free–clamped, each at the inner and outer radii. The various force or displacement conditions imposed on the inner and outer boundaries are realized by connecting the block networks according to the continuation conditions, as shown in Fig. 4; the restrained motion or the free motion is realized by opening or closing the corresponding port, respectively. The RF and the AF of the first three modes calculated by the impedance analysis and the FEM are listed in Table I, and they are in good agreement. Obviously, the free–clamped conditions give the higher RF and AF than the clamped–free conditions. Interestingly, the free–free conditions also give the higher frequencies than the clamped–free conditions.

In the second case, the effects of the size of the piezoelectric layers partially covering the shim layer on the electromechanical behavior of the PAB are investigated; the configuration is shown in Fig. 5(a). The whole PAB, as shown in Fig. 5(b), can be divided into three sections, i.e., the inner extended shim layer (Sec. I), the piezoelectric layers covering the shim layer (Sec. III), and the outer extended shim layer (Sec. IV). The impedance matrices of Secs. II and IV can be represented by the mechanical impedance $\mathbf{Z}_\mathrm{M}$ in Eq. (30a) with the corresponding values of the inner and outer radii and $h_p=0$. The impedance matrix of the whole system is then obtained by connecting in series the impedance matrices of the three sections, as shown in Fig. 5(c). Two cases of boundary conditions are studied; the free–clamped condition is realized by closing all ports except the seventh and eighth ports, and the free–free condition is realized by closing all ports. The values of the inner and outer radii of the middle shim layer are $r_a=10$ mm and $r_b=30$ mm, respectively. The radial width and radial center position of the piezoelectric layers are denoted by r_w and r_c, respectively.

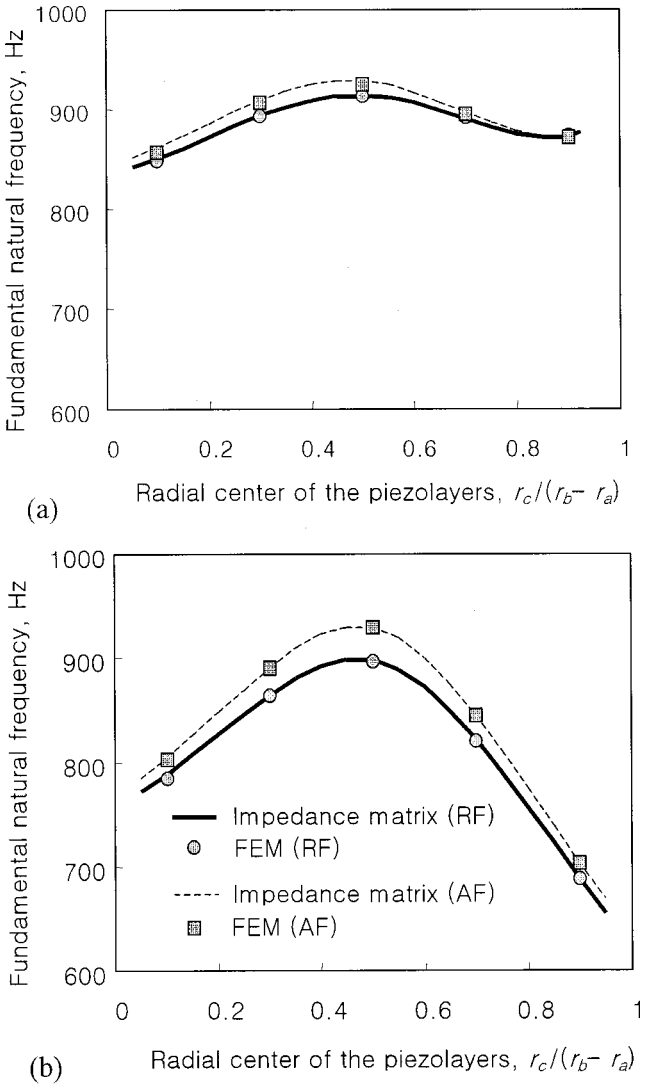


FIG. 6. The resonance frequencies (RF) and the antiresonance frequencies (AF) of the partially covered PAB versus the radial center of the piezoelectric layers r_c. The radial width $r_w/(r_b-r_a)=0.05$. (a) Free at inside and clamped at outside. (b) Free at inside and outside.

The RF and the AF of the fundamental mode for the various values of r_c with the radial width r_w kept equal to $0.05\times(r_b-r_a)$ are calculated by using the impedance matrix and the FEM under the two boundary conditions, each shown in Figs. 6(a) and (b). Notice that, in both free–clamped and free–free boundary conditions, the fundamental RF and AF reach their maximum values when the piezoelectric layers are located near the center of the shim layer. The fundamental RF and AF for the various values of r_w are also calculated by using the impedance matrix and the FEM; in this case, the radial center position of the piezoelectric layers $r_c=0.5\times(r_b-r_a)$. The results for the two boundary conditions are shown in Figs. 7(a) and (b), respectively. Under the free–clamped condition the RF and the AF increase as the radial width of the piezoelectric layers increases; especially

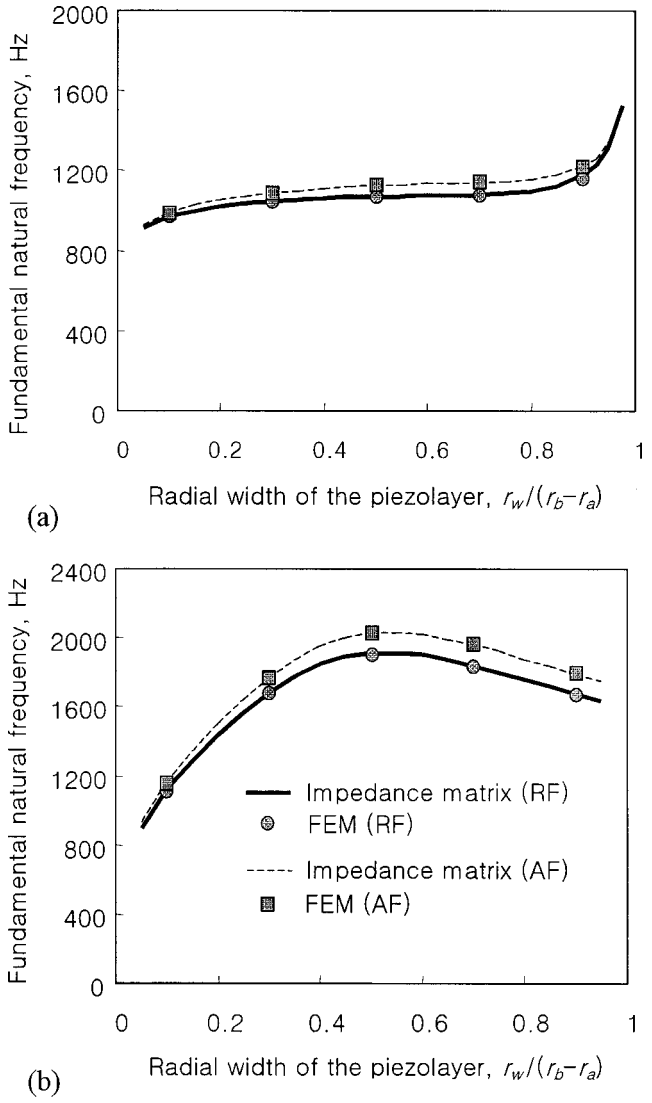


FIG. 7. The resonance frequencies (RF) and the antiresonance frequencies (AF) of the partially covered PAB versus the radial width of the piezoelectric layers r_w. The radial center $r_c/(r_b-r_a)=0.5$. (a) Free at inside and clamped at outside. (b) Free at inside and outside.

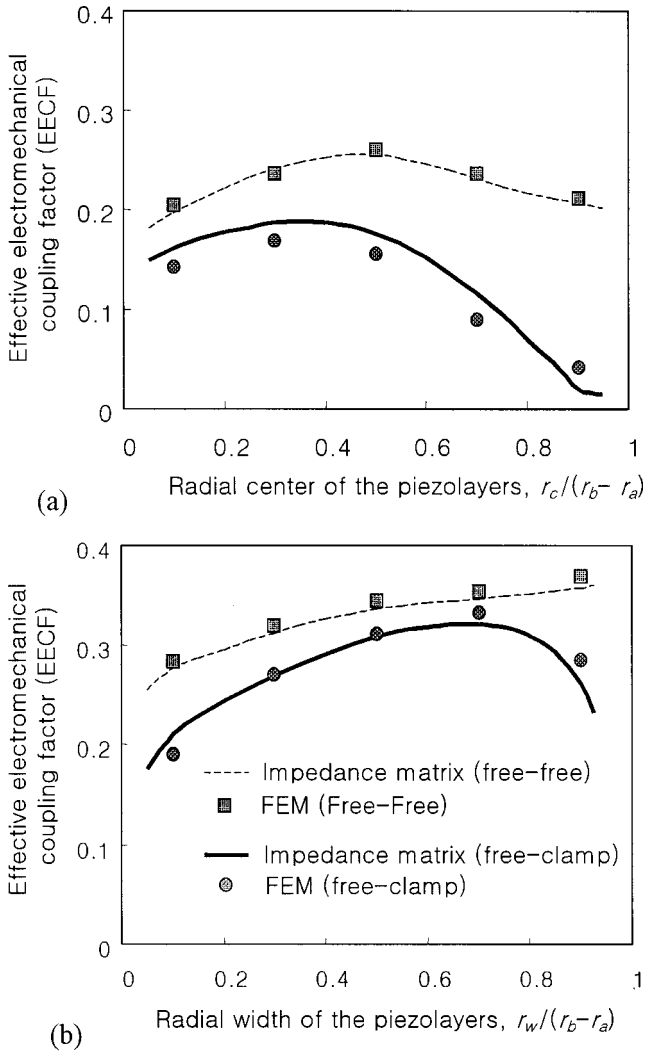


FIG. 8. The variations of the EECF of the partially covered PAB versus the radial width and the radial center. (a) The EEFC versus the radial center of the piezoelectric layers; the radial width $r_w/(r_b-r_a)=0.05$. (b) The EECF versus the radial width of the piezoelectric layers; the radial center $r_c/(r_b-r_a)=0.5$.

they sharply increase as the piezoelectric layers are covering the whole shim layer, i.e., $r_w/(r_b-r_a)\rightarrow 1$. In the case of the free–free condition, as the radial width of the piezoelectric layers increases, the RF and the AF increase until r_w reaches approximately $0.5\times(r_b-r_a)$ and then decrease afterward; they reach the maximum values when the piezoelectric layers are covering approximately half of the shim layer. Obviously, such a variation is closely related with the vibration mode shape. The results by the impedance matrix are in excellent agreement with those by the FEM. Obviously, as shown in Figs. 6 and 7, the AF are always greater than the RF in both cases; the differences are varying according to the geometric dimensions. The variations of the EECF under each boundary condition versus r_c and r_w are calculated using Eq. (33) and shown in Fig. 8, the EECF versus r_c in Fig. 8(a) and the EECF versus r_w in Fig. 8(b).

In the last case, the electromechanical behavior of the PCB with the piezoelectric layers partially covering the shim layer, of which configuration is shown in Fig. 9(a), is analyzed. The whole PCB, as shown in Fig. 9(b), can be divided into two sections, i.e., the piezoelectric layers covering the shim layer (Sec. II) and the outer extended shim (Sec. III). The impedance matrix of Sec. III can be represented by the mechanical impedance $\mathbf{Z}_{\mathrm{M}}$ in Eq. (30a). The impedance matrix of the whole system is then obtained by connecting in series the impedance matrices of the two sections, as shown in Fig. 9(c); at the first and second ports, the displacement and the rotation at the junction should be continuous. Two cases of boundary conditions at the outer radius are studied; the clamped conditions is realized by closing all ports except the third and fourth ports, and the free conditions is realized by closing all ports. The outer radius of the shim layer is kept equal to 30 mm. The variations of the RF and the AF versus the radius of the piezoelectric layers r_w are shown in Fig. 10 for the two boundary conditions. The results calculated by the impedance matrix are also in excellent agreement with those by the FEM.

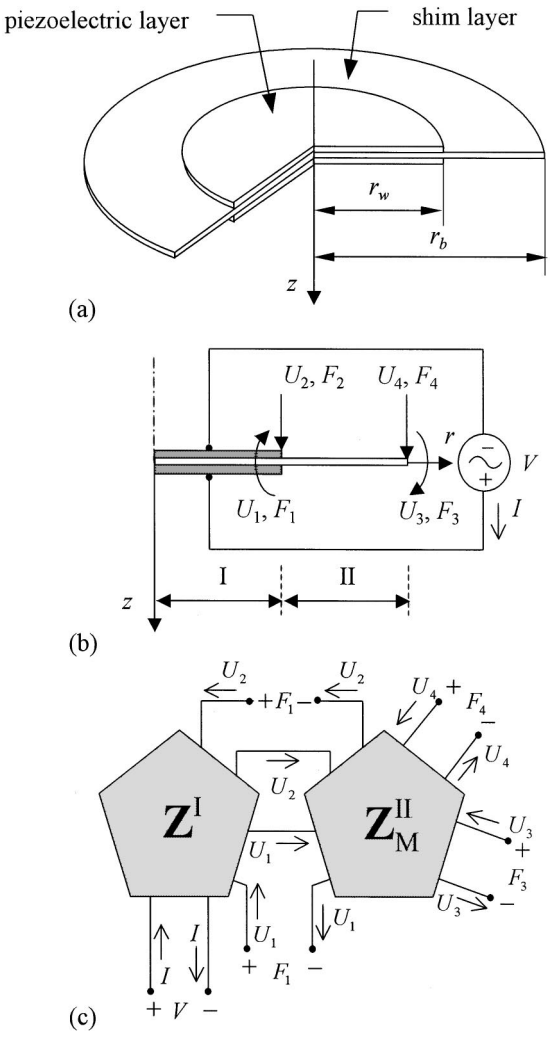


FIG. 9. (a) PCB with the piezoelectric layers partially covering the shim layer; (b) two sections with flow (U_i, I) and effort (F_i, V) vectors; and (c) each section represented by the impedance matrix, all in series connection.

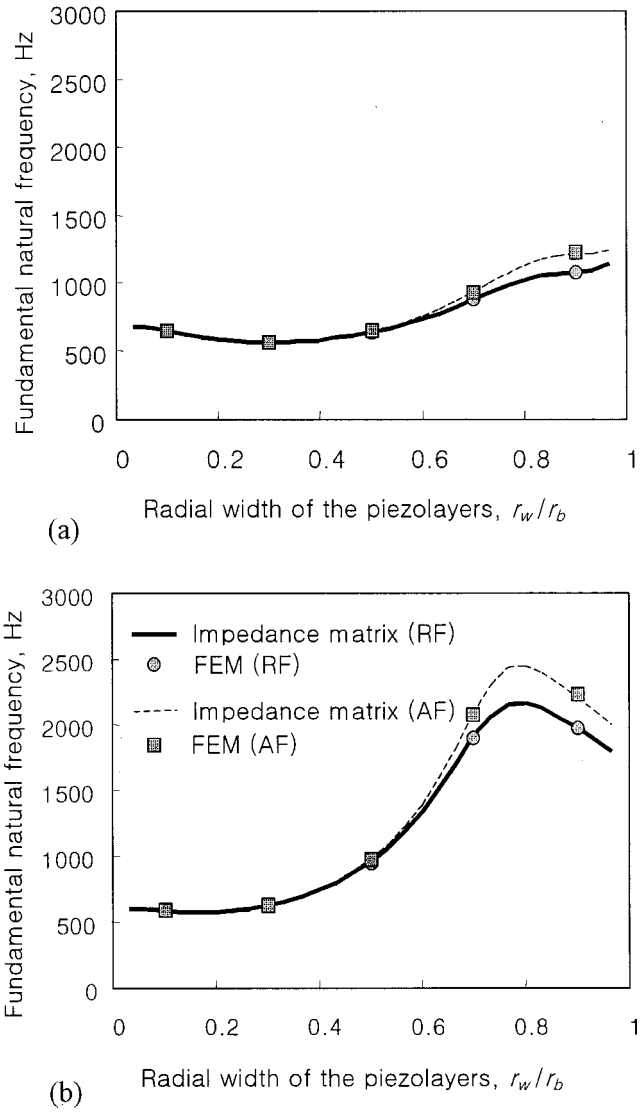


FIG. 10. The resonance frequencies (RF) and the antiresonance frequencies (AF) of the partially covered PAB versus the radial width of the piezoelectric layers r_w. (a) Clamped at outside and (b) free at outside.

V. CONCLUSIONS

The PAB and the PCB including those with the piezoelectric layers partially covering the shim layer can be now conveniently analyzed using the impedance and admittance matrices developed in this study, rather than using numerical methods such as the FEM.[7–9,18,25] It is shown that the block network representation together with matrices is very helpful in identifying the parameters of the annular or circular bimorphs, especially when the various force or displacement conditions are imposed on the inner and outer boundaries.

It is expected that using the impedance matrices the dimensions and locations of the multiple-rim-type piezoelectric layers which partially cover the shim layer can be optimally selected to assure the desired resonance and the antiresonance frequencies. In this case, the external mechanical or electrical components can be additionally connected in series or parallel to the PAB; this situation can be also analyzed using the impedance matrices and the block network presented in this article. It is also expected that the present analysis can be extended to include the harmonic extensional motions together with the flexural motion, which often occurs in the asymmetric bimorphs.

ACKNOWLEDGMENT

This research was supported by the Center of Innovative Design Optimization Technology (iDOT), Korea Science and Engineering Foundation.

[1] K. Uchino, *Piezoelectric Actuator and Ultrasonic Motors* (Kluwer Academic, Boston, 1996).

[2] K. Uchino, ''Recent topics of ceramic actuators: How to develop new ceramic devices,'' Ferroelectrics **91**, 281–292 (1989).

[3] P. Li and Y. Wen, ''Image resolution improvement by swing CCD array imager in two dimensions and its signal processing,'' Proc. SPIE **2308**, Part 2, 915–920 (1994).

[4] S. Akamine, T. R. Albrecht, M. J. Zdeblick, and C. F. Quate, ''A Planar process for Microfabrication of a Scanning Tunneling Microscope,'' Sens. Actuators A **23**, 964–970 (1990).

[5] J. Bost, ''A new type of tweeter horn employing a piezoelectric driver,'' J. Audio Eng. Soc. **23**, 796–801 (1978).

[6] J. Bost, ''A new piezoelectric driver enhances horn performance,'' J. Audio Eng. Soc. **28**, 244–249 (1980).

[7] G. Perçin, L. Levin, and B. T. Khuri-Yakub, ''Piezoelectrically actuated droplet ejector,'' Rev. Sci. Instrum. **68**, 4561–4563 (1997).

[8] G. Perçin, A. Atalar, F. L. Degertekin, and B. T. Khuri-Yakub, ''Micro-

machined two-dimensional array piezoelectrically actuated transducers,'' Appl. Phys. Lett. **72**, 1397–1399 (1998).
[9] G. Perçin, T. S. Lundgren, and B. T. Khuri-Yakub, ''Controlled ink-jet printing and deposition of organic polymers and solid particles,'' Appl. Phys. Lett. **73**, 2375–2377 (1998).
[10] J. G. Smits, S. I. Dalke, and T. K. Cooney, ''The Constitutent Equations of Piezoelectric Bimorphs,'' Sens. Actuators A **28**, 41–61 (1991).
[11] J. G. Smits and A. Ballato, ''Dynamic Admittance Matrix of Piezoelectric Cantilever Bimorphs,'' J. Microelectromech. Syst. **3**, 105–112 (1994).
[12] J. G. Smits, W. Choi, and A. Ballato, ''Resonance and Antiresonance of Symmetric and Asymmetric Cantilevered Piezoelectric Flexors,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **44**(2), 250–258 (1997).
[13] H. Tanaka, ''Generalized basic equations for bending motions of piezoelectric bars from Hamilton's principle,'' J. Acoust. Soc. Am. **95**, 1768–1772 (1994).
[14] H. Tanaka and R. Aoyagi, ''Derivation of Equivalent Circuit of Multilayered Bending Vibration,'' in Proceedings of Acoustic Engineering, Tohoku Univ., Japan, February 1992, pp. 262–263 (in Japanese).
[15] R. Aoyagi and H. Tanaka, ''Equivalent Circuit Analysis of Piezoelectric Bending Vibrators,'' Jpn. J. Appl. Phys. **33**(5B), Part 1, 3010–3014 (1995).
[16] Q. M. Wang, X. H. Du, B. Xu, and L. Eric Cross, ''Electromechanical Coupling and Output Efficiency of Piezoelectric Bending Actuator,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **46**(3), 638–646 (1999).
[17] Y. S. Cho, Y. E. Park, C. S. Han, and S. K. Ha, ''Five-port Equivalent Circuit of Piezoelectric Bimorph Beam,'' Sens. Actuators A **84**, 140–148 (2000).
[18] A. B. Dobrucki and P. Pruchnicki, ''Theory of piezoelectric bimorph,'' Sens. Actuators A **58**, 203–212 (1997).
[19] H. L. Hui, ''Axisymmetric response of circular plates with piezoelectric layers: An exact solution,'' Int. J. Mech. Sci. **40**(12), 1265–1279 (1998).
[20] P. C. Y. Lee and J. D. Yu, ''Governing Equations for a Piezoelectric Plate with Graded Properties Across the Thickness,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **45**(1), 236–250 (1998).
[21] P. C. Y. Lee, J. D. Yu, X. Li, and W. H. Shih, ''Piezoelectric Ceramic Disks with Thickness-Graded Material Properties,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **46**(1), 205–215 (1999).
[22] B. A. Auld, *Acoustic Fields and Waves in Solids*, 2nd ed. (Krieger, Melbourne, 1990).
[23] A. W. Leissa, *Vibration of Plates*, NTIS, NASA SP-160, N70-18461.
[24] *IEEE Standard on Piezoelectricity*, ANSI/IEEE Std. 176, 1987, p. 43.
[25] S. K. Ha, C. Keilers, and F. K. Chang, ''Finite Element Analysis of Composite Structures Containing Distributed Piezoceramic Sensors and Actuators,'' AIAA J. **30**(3), 772–780 (1992).

Measured capacitance of a condenser microphone as a function of diaphragm displacement

Xinche Yan[a] and Malcolm J. Crocker
Department of Mechanical Engineering, Auburn University, Alabama 36849

Li Jun Zeng
TRW Automotive, 12025 Tech Center Drive, Livonia, Michigan 48150

(Received 3 January 2000; revised 20 May 2000; accepted 29 July 2000)

This paper shows that the capacitance of a polarized condenser microphone measured with a sinusoidal electrical signal is quantitatively related to the displacement of the microphone diaphragm. In fact, the capacitance measured with an ac signal is the ''input capacitance'' of the microphone when it acts as a transmitter. The expressions for the diaphragm displacement and the input capacitance are derived with damping taken into account. The input capacitance differs from the static capacitance with bias by an additional term which varies with the polarization voltage and frequency. Therefore, in general, it is impossible to measure the static capacitance with bias using a low frequency ac signal without the diaphragm being blocked. In order to measure the true static capacitance with bias, a static method based on the dc voltage distribution is presented. The experimental results for the input capacitance and the static capacitance with bias, obtained with the abovementioned ac and dc methods, respectively, give a high level of confidence in the theory. The experimental results reported by some others also support the theory. © *2000 Acoustical Society of America.* [S0001-4966(00)00611-1]

PACS numbers: 43.38.Bs, 43.40.At, 43.40.Dx [SLE]

LIST OF SYMBOLS

a	radius of diaphragm and backplate (m)
B_1, B_n	dimensionless correction factor due to negative stiffness, $-S_d$
C	instantaneous capacitance (F)
C_{00}, C_0	static capacitance without bias and with bias (F)
$\mathbf{C}_1$	complex magnitude of incremental capacitance due to diaphragm motion (F)
C_{in}	input capacitance, i.e., the capacitance derived from the input susceptance of a condenser microphone when it acts as a transmitter (F)
C_s, C_c, C_w	values of capacitance standards and stray capacitances (F)
c	velocity of transverse waves in diaphragm (m/s)
c_1	dimensionless coefficient of series expression for $\mathbf{k}a$
$F_C(f), F_G(f)$	frequency factors
f	frequency (Hz)
f_1, f_2, f_n	natural frequencies for symmetric modes of diaphragm vibration (Hz)
f_a, f_b	frequencies for maximum and minimum values of C_{in} beside f_1 (Hz)
G	input conductance ($S=\Omega^{-1}$)
h	static distance between diaphragm and backplate with axis of condenser microphone aligned horizontally without any voltage applied (m)
i	alternating current flowing through condenser microphone (A)
$J_0(z), J_1(z), J_2(z)$	Bessel functions of order zero, one, and two with complex argument
j	$(-1)^{1/2}$
j_{01}, j_{0n}	roots of equation $J_0(z)=0$ ($j_{01}=2.4048$)
K	stiffness of boundary per unit perimeter length (N/m^2)
$k_d, k_s, \mathbf{k}$	wave numbers (m^{-1})
M	mass of boundary per unit perimeter length (kg/m)
n	integer
P	air pressure in vacuum chamber (mm Hg=Torr)
P_d, P_s	electrostatic attraction forces acting on diaphragm per unit area (Pa)
q	instantaneous electric charge on diaphragm (C)
R	mechanical resistance of boundary per unit perimeter length (Ns/m^2)
R_0	Insulation resistance between electrodes (Ω)
R_{in}	input resistance (Ω)
r	radial coordinate of diaphragm (m)
S_d, S_s	absolute values of negative stiffness of diaphragm per unit area (N/m^3)
T	tension in diaphragm (N/m)
$u, \mathbf{u}_A$	velocity of diaphragm and its complex amplitude (m/s)
t	time (s)
V_0	polarization voltage (V)
V_a, V_i	direct-current voltage (V)

[a]Present address: Knowles Electronics, LLC, 1151 Maplewood Drive, Itasca, IL 60143.

V_{clps}	minimum polarization voltage causing diaphragm collapse (V)
v_0	amplitude of sinusoidal excitation voltage across condenser microphone (V)
$X(r)$	space factor of dynamic displacement of diaphragm x_d (m)
x	displacement of diaphragm (m)
x_d, $\mathbf{x_{dA}}$	dynamic displacement of diaphragm and its complex amplitude (m)
x_s	static displacement of diaphragm (m)
Y	input electric admittance ($S=\Omega^{-1}$)
Z	mechanical impedance of boundary per unit perimeter length (Ns/m^2)
$\alpha, \alpha_1, \alpha_n$	attenuation coefficients (s^{-1})
β	viscoelastic material constant (s)
δ	complex, dimensionless viscous boundary damping parameter ($\delta=\eta+j\xi$)
ϵ_0	permittivity of free space (8.854×10^{-12} F/m)
θ, θ_n	phase angles (rad)
π	ratio of circumference to diameter of circle
σ	surface mass density of diaphragm (kg/m^2)
ω	angular frequency (rad/s)
$\boldsymbol{\omega_d}, \omega_d, \omega_{dn}$	natural angular frequencies for damped, free diaphragm vibration (rad/s)
ω_1, ω_n	natural angular frequencies for symmetric modes of diaphragm vibration (rad/s)

I. INTRODUCTION

Although the capacitance of an electrostatic transducer is a well-known physical quantity in electroacoustics, the difference between the blocked capacitance and the free capacitance of condenser microphones, as described in textbooks and the American National Standard,[1–5] has not attracted enough attention from researchers. For example, some have tried to measure the static capacitance of polarized electrostatic earphones or condenser microphones and have obtained noticeable systematic errors because the motion of the transducer diaphragm has been ignored.[6–8] Thus what they have actually measured is approximately the free capacitance, as pointed out by Beranek[2] and Hunt.[3] In fact, the unblocked diaphragms of the transducers being measured should be excited into vibration by ac test signals. Hence the measured results are no longer the static capacitances. Even though the displacements of the diaphragms are very small, their effects cannot be neglected except in some special cases. The purpose of this paper is to show that, in general, the capacitance of a polarized condenser microphone measured with a sinusoidal electrical signal is related to the diaphragm displacement, and that, to measure the true static capacitance of a polarized condenser microphone without the effect of diaphragm vibration, a dc method should be used. Such a quantitative relationship and a corresponding static method have not been reported in the literature so far.

For the capacitance of a condenser microphone, at least the following different concepts are applicable: (1) the static capacitance without bias, C_{00}—an imaginary blocked capacitance with no diaphragm deflection; (2) the static capacitance with bias, C_0—also an imaginary blocked capacitance but with a static diaphragm deflection due to the polarization voltage; (3) the instantaneous capacitance, C—a time varying ''capacitance'' including a steady term, C_0, and a complex incremental part due to the dynamic diaphragm motion; and (4) the capacitance of a microphone when it acts as a transmitter,[9] C_{in}. Due to its reciprocal properties, a condenser microphone may operate not only as a receiver but also as a transmitter. As shown in Fig. 1, the diaphragm moves in response to an alternating voltage applied across the electrodes, and radiates a corresponding acoustic wave if the transducer is placed in air. At the same time, the microphone presents an electric impedance or admittance to the electric circuit to which it is connected. The capacitance C_{in} derived from the input susceptance of the microphone, and called the ''input capacitance'' for short, forms the fundamental theoretical concept of this paper. This concept has been employed previously in the pressure (coupler) calibration procedure of condenser microphones.[9] The calibration technique uses the reciprocity method in which two microphones are mounted on a coupler, one acting as a transmitter and the other as a receiver. The value of the input capacitance depends on the load on the microphone diaphragm. If there is no force acting on the diaphragm, the input capacitance is equal to the free capacitance; if the diaphragm is restrained from movement, the input capacitance becomes the blocked capacitance.

The problem of measuring the capacitance of condenser microphones arises from the property of the test signals. Currently no methods exist for measuring the capacitance of condenser microphones without the application of ac signals to the microphones. It is suggested in the relevant American National Standard[10] that the electrical driving-point impedance of a condenser microphone can be measured by means of a suitable bridge, such as the Wagner-grounded Schering

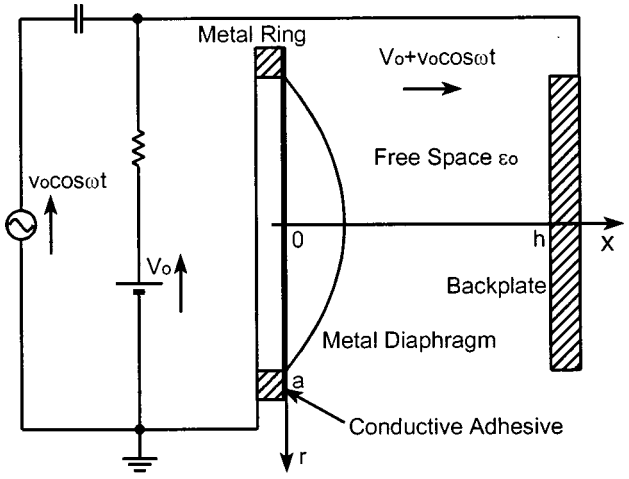


FIG. 1. Schematic for a condenser microphone when it acts as a transmitter.

bridge, when the impedance is mainly capacitive. It is common knowledge that capacitance bridges are ac measurement devices. Besides, the well-known condenser microphone manufacturers have measured the capacitance of their products as a function of frequency or at a certain chosen frequency.[11] This implies that ac test signals must have been used in such measurements. When a condenser microphone is fed with ac test signals, it must act as a transmitter to a certain extent as long as its diaphragm is not blocked and a polarization voltage is applied. This is the situation most diaphragms are faced with, because these diaphragms are too thin and frail to be blocked by any solid objects. In the case of the pressure (coupler) calibration, even the diaphragm is loaded by an acoustic impedance, the diaphragm is still not blocked.[9,10] Thus the capacitance of a polarized condenser microphone measured with an ac signal is, in fact, the input capacitance. In general, the static capacitance with bias cannot be measured using low frequency ac signals without the diaphragm being blocked.

In both the theory and the experiments reported in this study, the condenser microphone cartridge used was larger than usual and was designed so that: (1) the diameters of the diaphragm and the backplate are the same; (2) the backplate is solid and flat; and (3) the air in the gap between the electrodes can be pumped out through some holes in the insulator for holding the backplate. The reasons for such a design are: (1) to simplify the equation for the diaphragm motion and the expressions for the microphone capacitance; (2) to make the diaphragm displacement and the microphone capacitance larger, so that the measurements on the capacitance will be more nearly accurate; and (3) to simplify the manufacture of the microphone. Since the derivation for the input capacitance is independent of the construction of the microphone used, the dependence of the input capacitance on the diaphragm displacement is universally applicable.

II. THEORY

A. Microphone capacitance

The capacitance of the microphone is proportional to the area of the electrodes and inversely proportional to the distance between them. The differences between C_{00}, C_0, and C are due to the different gaps between the electrodes, i.e., due to the differences in diaphragm displacements. Since the microphone is situated in a vacuum chamber, there is no acoustic excitation causing any diaphragm displacement. When no excitation voltage is applied across the electrodes, after the free vibration has decayed away, the diaphragm can only have a static displacement caused by the polarization voltage. When[12]

$$V_0 \ll V_{\mathrm{clps}} = \sqrt{\frac{8}{5.064}}\sqrt{\frac{Th^3}{\epsilon_0 a^2}}, \tag{1}$$

no collapse will occur, and $x_s \ll h$. Collapse means that the diaphragm is being attracted sufficiently to make contact with the backplate. When $\nu_0 \ll V_0$, the diaphragm displacement may be divided into two parts:[8,12] $x = x_s + x_d$, where $x_d \ll x_s$ and $x_d = |x_d| \exp[j(\omega t - \theta)] = \mathbf{x_{dA}} \exp(j\omega t)$ as is shown later. Then the following expressions may be written:

$$C_{00} = \frac{\epsilon_0 \pi a^2}{h}, \tag{2}$$

$$C_0 = \epsilon_0 \int_0^a \frac{2\pi r\, dr}{h - x_s} = \frac{\epsilon_0}{h}\int_0^a 2\pi \left(1 - \frac{x_s}{h}\right)^{-1} r\, dr \approx C_{00}\left(1 + \frac{\langle x_s \rangle}{h}\right), \tag{3}$$

and

$$C = \epsilon_0 \int_0^a \frac{2\pi r\, dr}{h - x} = \epsilon_0 \int_0^a \frac{2\pi r}{h - x_s}\left(1 - \frac{x_d}{h - x_s}\right)^{-1} dr \approx C_0 + \mathbf{C_1} e^{j\omega t}, \tag{4}$$

where

$$\langle x_s \rangle = \frac{1}{\pi a^2}\int_0^a x_s \cdot 2\pi r\, dr, \tag{5}$$

and

$$\mathbf{C_1} = 2\pi\epsilon_0 \int_0^a \frac{\mathbf{x_{dA}} r\, dr}{(h - x_s)^2}. \tag{6}$$

The expression for C_0 is equivalent to that derived by Hawley *et al.*[13] since the latter can be obtained by integration using Eq. (3) and then making an approximation.

Finally, C_{in} is derived from the input electric admittance as follows. First,

$$q = (V_0 + \nu_0 e^{j\omega t})C \approx (V_0 + \nu_0 e^{j\omega t})(C_0 + \mathbf{C_1} e^{j\omega t}), \tag{7}$$

where the applied sinusoidal potential difference $\nu_0 \cos \omega t$ has been replaced by $\nu_0 \exp(j\omega t)$ to simplify the mathematical manipulation. Then the resultant current at the same point is given by

$$i = \frac{dq}{dt} \approx j\omega[\nu_0 C_0 + V_0 \mathbf{C_1}] e^{j\omega t}. \tag{8}$$

The input electric admittance is obtained by taking the reciprocal of the driving-point impedance:[5,10]

$$Y = \frac{i}{\nu_0 e^{j\omega t}} = j\omega C_{\mathrm{in}} + G, \tag{9}$$

where $j\omega C_{\mathrm{in}}$ is the input susceptance,

$$C_{\mathrm{in}} = C_0 + \frac{V_0}{\nu_0}\mathrm{Re}(\mathbf{C_1}) = C_0 + \frac{V_0}{\nu_0}\cdot 2\pi\epsilon_0 \int_0^a \frac{\mathrm{Re}(\mathbf{x_{dA}}) r\, dr}{(h - x_s)^2}, \tag{10}$$

and

$$G = j\omega \frac{V_0}{\nu_0} j \cdot \mathrm{Im}(\mathbf{C_1}) = \frac{V_0}{\nu_0}\cdot 2\pi\epsilon_0 \int_0^a \frac{\mathrm{Re}(\mathbf{u_A}) r\, dr}{(h - x_s)^2}. \tag{11}$$

Equation (9) indicates that, in an ac circuit, the microphone may be considered as consisting of a capacitor C_{in} and

a resistor $R_{\text{in}}=1/G$ connected in parallel. The resistance is caused by $\text{Im}(\mathbf{C_1})$ or $\text{Re}(\mathbf{u_A})$, and represents the energy loss due to damping, while the second term of the capacitance is caused by $\text{Re}(\mathbf{x_{dA}})$. Besides, a high resistance R_0 should be included in G. Since R_0 and R_{in} are in parallel, $G=(1/R_{\text{in}})+(1/R_0)$. For frequencies close to the first resonance of the diaphragm, where the polarization voltage varied from 10 to 100 V, the theoretical value of R_{in} is $10^8-10^{10}\,\Omega$, and the experimental value is $10^7\,\Omega$. They are much lower values than that of R_0; thus R_0 can be neglected.

B. Diaphragm displacement and velocity

In this study, only the diaphragm motion with circular symmetry was considered. This is because, for all the asymmetric modes, the average displacement over the diaphragm is zero[1] and makes no contribution to any changes in microphone capacitance. Thus there is only one spatial variable in the partial differential equation governing the diaphragm motion:

$$T\frac{\partial^2 x}{\partial r^2}+\frac{T}{r}\frac{\partial x}{\partial r}+\beta T\frac{\partial^3 x}{\partial r^2\,\partial t}+\beta\frac{T}{r}\frac{\partial^2 x}{\partial r\,\partial t}+\frac{\epsilon_0(V_0+\nu_0\cos\omega t)^2}{2(h-x)^2}=\sigma\frac{\partial^2 x}{\partial t^2}, \tag{12}$$

where all the terms on the left-hand side are forces per unit area acting on the diaphragm. The first two terms represent the elastic restoring force due to tension; the third and fourth terms are the resistance forces caused by the material damping; and the fifth term is the electrostatic attraction force. The effect of gravity on the diaphragm displacement has been eliminated by aligning the microphone axis horizontally. Since the surface mass density of the diaphragm is very small, it can be shown that the additional diaphragm tension caused by the gravitational force can be neglected when the axis of the microphone is aligned horizontally. There are no air reaction (loading) forces because the microphone is assumed to be situated in a vacuum chamber. The purpose of placing the microphone in a vacuum chamber is to simplify the equation of motion and allow a larger diaphragm displacement. This is because the aim of this study is to find out the relationship between the input capacitance and the diaphragm displacement rather than the air reaction.

There are at least two possible sources of damping to be taken into account. First, consider the internal damping. Assume that the diaphragm material, i.e., stainless steel, is a Voigt viscoelastic solid. For such a solid, there are two terms in the stress–strain law, where one of them represents the viscoelastic damping.[14] Since the diaphragm tension is the product of the stress and the thickness, the tension also includes a damping term. That is why, in Eq. (12), the third and fourth terms are included. Next, consider the damping which comes from the boundary of the diaphragm. This damping does not affect Eq. (12). Hence it is only included in the boundary condition:[1]

$$-T\left.\frac{\partial x}{\partial r}\right|_{r=a}=R\left.\frac{\partial x}{\partial t}\right|_{r=a}+M\left.\frac{\partial^2 x}{\partial t^2}\right|_{r=a}+Kx|_{r=a}=\left[R+j\left(\omega M-\frac{K}{\omega}\right)\right]\left.\frac{\partial x}{\partial t}\right|_{r=a}=Z\left.\frac{\partial x}{\partial t}\right|_{r=a}, \tag{13}$$

where the right-hand side is expressed only in terms of the velocity, because the steady-state diaphragm motion is a simple harmonic function of time when the excitation voltage is sinusoidal.

In this study, $\nu_0=(0.001)^{1/2}\ \text{V}=31.6\,\text{mV}$, $V_0=10-100\,\text{V}$, and V_{clps} is about 1000 V, i.e., $\nu_0\ll V_0\ll V_{\text{clps}}$. Therefore, $x=x_s+x_d$, where $x_s, x_d\ll h$. By substituting $\nu_0=0$ and $x=x_s$ into Eq. (12), the linearized equation for the static displacement is obtained:

$$\frac{d^2x_s}{dr^2}+\frac{1}{r}\frac{dx_s}{dr}+\frac{S_s}{T}x_s=-\frac{P_s}{T}, \tag{14}$$

where

$$S_s=\frac{\epsilon_0 V_0^2}{h^3}, \tag{15}$$

and

$$P_s=\frac{\epsilon_0 V_0^2}{2h^2}. \tag{16}$$

By substituting $x=x_s$ into Eq. (13), the boundary condition may be written as

$$x_s|_{r=a}=-\frac{T}{K}\left.\frac{dx_s}{dr}\right|_{r=a}. \tag{17}$$

The solution to Eq. (14) which satisfies Eq. (17) is given by

$$x_s=\frac{h}{2}\left(\frac{J_0(k_s r)}{J_0(k_s a)-\dfrac{Tk_sJ_1(k_sa)}{K}}-1\right), \tag{18}$$

where

$$k_s=\sqrt{S_s/T}. \tag{19}$$

Because the ratio T/K is much smaller than unity when $T=1967\,\text{N/m}$ and K is of the order of magnitude of $10^7\,\text{N/m}^2$, and k_s is less than $2.5\,\text{m}^{-1}$, Eq. (18) can be written as

$$x_s\approx\frac{h}{2}\left(\frac{J_0(k_s r)}{J_0(k_s a)}-1\right). \tag{20}$$

Thus

$$x_s|_{r=a}\approx 0. \tag{21}$$

Substituting x_s+x_d for x into Eqs. (12) and (13), then using Eqs. (14) and (21), yields the linearized equation and the boundary condition for the dynamic displacement:

$$T\frac{\partial^2 x_d}{\partial r^2}+\frac{T}{r}\frac{\partial x_d}{\partial r}+\beta T\frac{\partial^3 x_d}{\partial r^2\partial t}+\beta\frac{T}{r}\frac{\partial^2 x_d}{\partial r\,\partial t}-\sigma\frac{\partial^2 x_d}{\partial t^2}+S_d x_d=-P_d e^{j\omega t}, \tag{22}$$

and

$$x_d|_{r=a} = -\frac{T}{j\omega Z}\left.\frac{\partial x_d}{\partial r}\right|_{r=a}, \tag{23}$$

where

$$S_d = \frac{\epsilon_0 V_0^2}{(h-\langle x_s\rangle)^3}, \tag{24}$$

and

$$P_d = \frac{\epsilon_0 V_0 \nu_0}{(h-\langle x_s\rangle)^2}. \tag{25}$$

In Eq. (22), for the convenience of manipulation, $\cos \omega t$ and $h-x_s$ have been replaced by $\exp(j\omega t)$ and $h-\langle x_s\rangle$, respectively, where

$$\langle x_s\rangle = \frac{1}{\pi a^2}\int_0^a x_s \cdot 2\pi r\, dr \approx \frac{h}{2}\frac{J_2(k_s a)}{J_0(k_s a)}. \tag{26}$$

When the electric excitation is simple harmonic, the steady-state mechanical response should also be simple harmonic: $x_d = X(r)\exp(j\omega t)$. Therefore, Eqs. (22) and (23) may be changed into the following forms:

$$\frac{d^2X}{dr^2} + \frac{1}{r}\frac{dX}{dr} + k_d^2 X = -\frac{P_d}{(1+j\omega\beta)T}, \tag{27}$$

and

$$X|_{r=a} = -\frac{T}{j\omega Z}\left.\frac{dX}{dr}\right|_{r=a}, \tag{28}$$

where

$$k_d = \sqrt{\frac{\omega^2\sigma + S_d}{(1+j\omega\beta)T}}. \tag{29}$$

By solving for $X(r)$ the steady-state dynamic displacement and the velocity may be written as

$$x_d = \frac{P_d}{\omega^2\sigma + S_d}\left[\frac{J_0(k_d r)}{J_0(k_d a) - \dfrac{T k_d J_1(k_d a)}{j\omega Z}} - 1\right] e^{i\omega t}$$
$$= |x_d| e^{-j\theta} e^{j\omega t} = \mathbf{x_{dA}} e^{j\omega t}, \tag{30}$$

and

$$u = j\omega |x_d| e^{-j\theta} e^{j\omega t} = j\omega \mathbf{x_{dA}} e^{j\omega t} = \mathbf{u_A} e^{j\omega t}. \tag{31}$$

C. Resonance and damping

The values of the unknowns β and Z may be found through attenuation since all of them involve energy loss. The attenuation is usually studied in the neighborhood of resonances. To find the natural angular frequencies for the vibration expressed by Eq. (30), the following two extreme cases are discussed.

First, assume that all the damping is contributed by the diaphragm material, and the boundary is rigid without any damping, i.e., $R=0$, $K\to\infty$, $Z\to\infty$, and $x_d|_{r=a}=0$, then there are two equivalent expressions for x_d. One is directly derived from Eq. (30):

$$x_d = \frac{P_d}{\omega^2\sigma + S_d}\left[\frac{J_0(k_d r)}{J_0(k_d a)} - 1\right] e^{j\omega t}, \tag{32}$$

and the other is obtained using the normal-mode summation:[15]

$$x_d = 2\frac{P_d}{\sigma}\sum_{n=1}^{\infty} J_0\left(\frac{j_{0n}}{a} r\right) \times \frac{1}{j_{0n} J_1(j_{0n})}\frac{e^{j(\omega t-\theta_n)}}{\sqrt{(\omega_n^2-\omega^2)^2 + (2\alpha_n\omega)^2}}, \tag{33}$$

where

$$\theta_n = \arctan\frac{2\alpha_n\omega}{\omega_n^2-\omega^2}, \tag{34}$$

$$\omega_n = 2\pi f_n = \frac{j_{0n}}{a} c B_n, \tag{35}$$

$$\alpha_n = \frac{\beta}{2}\frac{j_{0n}^2}{a^2}c^2, \tag{36}$$

$$c = \sqrt{T/\sigma}, \tag{37}$$

and

$$B_n = \sqrt{1 - \frac{a^2 S_d}{j_{0n}^2 T}}. \tag{38}$$

Since Eqs. (32) and (33) should give exactly the same solution for x_d, then x_d in Eq. (32) must have the same frequency dependence as is clearly shown in Eq. (33).

Second, assume that all the damping occurs at the boundary, and the diaphragm itself is without damping, i.e., $\beta=0$. There are no loss terms in Eq. (22), but Eq. (23) is a damped boundary condition. The damping of the forced vibration governed by these two equations can also be described by an attenuation coefficient which should be the same as that in the damped, free vibration governed by

$$T\frac{\partial^2 x_d}{\partial r^2} + \frac{T}{r}\frac{\partial x_d}{\partial r} - \sigma\frac{\partial^2 x_d}{\partial t^2} + S_d x_d = 0 \tag{39}$$

and Eq. (23). To describe such a decaying vibration, a complex angular frequency $\boldsymbol{\omega_d} = \omega_d + j\alpha$ is needed.[1] By substituting $x_d = X(r)\exp(j\boldsymbol{\omega_d} t)$ into Eq. (39), $X(r) = J_0(\mathbf{k}r)$ is found, where

$$\mathbf{k} = \sqrt{\frac{\boldsymbol{\omega_d}^2\sigma + S_d}{T}} \tag{40}$$

is also complex. The solution for x_d should satisfy Eq. (23), where ω has now been replaced by $\boldsymbol{\omega_d}$:

$$J_0(\mathbf{k}a) - j\,\delta J_0'(\mathbf{k}a) = 0, \tag{41}$$

where

$$\delta = \eta + j\xi = \frac{\mathbf{k}T}{\boldsymbol{\omega}_\mathbf{d} Z}. \tag{42}$$

The quantities ω_d and α can be found using Chen's method.[16] When $\delta = 0$ (rigid boundary), $\mathbf{k}a = j_{0n}$, and $\mathbf{k}$ is real. When $|\delta| \ll 1$, $\mathbf{k}a$ is analytic with respect to δ in a small region close to each value of j_{0n} where $\mathbf{k}a = j_{0n} + c_1\delta + \cdots$ is a power series in δ. By substituting the series expression for $\mathbf{k}a$ into Eq. (41) and using the Taylor expansion, $c_1 = j$ can be found. Hence

$$\mathbf{k} = \frac{j_{0n}}{a} + j\frac{\delta}{a} = \frac{j_{0n} - \xi}{a} + j\frac{\eta}{a}. \tag{43}$$

Then, using Eq. (40), the results are obtained:

$$\omega_{dn} \approx (j_{0n} - \xi)\frac{c}{a} B_n, \tag{44}$$

and

$$\alpha_n \approx \eta \frac{c}{a} B_n^{-1}. \tag{45}$$

Because, in this study, B_n is close to 1,

$$\boldsymbol{\omega}_\mathbf{d} \approx [(j_{0n} - \xi) + j\,\eta B_n^{-2}]\frac{c}{a} B_n \approx \mathbf{k}cB_n. \tag{46}$$

Now return to the original forced vibration governed by Eqs. (22) and (23). Its attenuation coefficient is the same as α_n as expressed in Eq. (45), and its angular resonance frequency, ω_n, is approximately equal to ω_{dn} as shown in Eq. (44), since the damping is very light.

With ω_1 and α_1 measured, β and Z may be calculated using Eqs. (36) and (42), respectively, for the two extreme cases:

(1) Assuming β is the only source of damping,

$$\beta = 2\alpha_1 \frac{a^2}{j_{01}^2 c^2}. \tag{47}$$

(2) Assuming Z or R is the only source of damping,

$$Z = \frac{T}{cB_1\delta} = \frac{T}{cB_1}\left[\alpha_1 \frac{aB_1}{c} + j\left(j_{01} - \omega_1 \frac{a}{cB_1}\right)\right]^{-1}. \tag{48}$$

Because the value of Z, calculated from the measured data for ω_1 and α_1, is very large and almost a pure real quantity,

$$R = \mathrm{Re}(Z) \approx \frac{T}{aB_1^2\alpha_1}. \tag{49}$$

It is not surprising that α_1 is inversely proportional to R, since an increase in R simply causes the boundary to become closer to the infinite impedance condition for which the resonance peak approaches infinity. Therefore, when R is increased, the resonance peak becomes sharper, and α_1 is smaller.

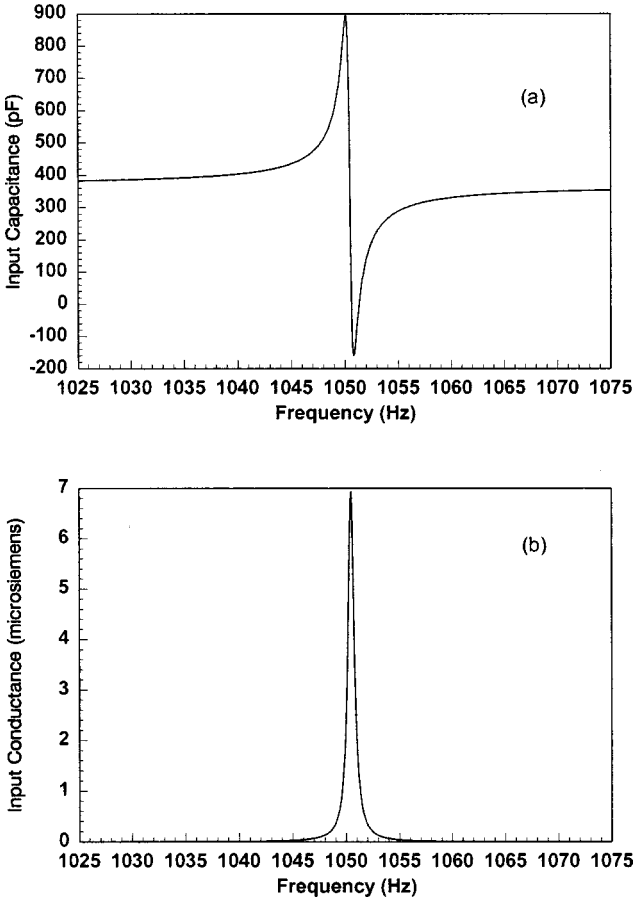


FIG. 2. (a) Input capacitance and (b) input conductance as a function of frequency showing the coincidence of three curves calculated using different expressions for x_d, Eqs. (30), (32), and (33), with $V_0 = 100\,\mathrm{V}$ and $\alpha_1 = 2.17\,\mathrm{s}^{-1}$.

D. Input capacitance and input conductance

Now x_d can be calculated using Eq. (32) (which contains Bessel functions with a complex argument due to β being nonzero) or Eq. (33) (a summation of normal modes) for Case 1, and using Eq. (30) (which contains Bessel functions with real arguments since $\beta = 0$) for Case 2. Then, C_{in} and G can be plotted as a function of f using the three expressions for x_d, respectively, with different values of V_0 and α_1. There are no noticeable differences found between these curves at frequencies close to and below f_1. Due to space limitations, only Figs. 2(a) and (b) for the frequencies near f_1 are shown. It seems that there is only one curve in each figure. In fact, there are three curves coincident with each other. As a function of r and f, x_d is not sensitive to the source of the damping, at least, within the scope of this study. The difficulty of identifying the contribution of the sources of the damping is thus avoided. It is possible to assume that all the damping is contributed by the diaphragm material, solely for ease of mathematical manipulation. Therefore, the normal-mode summation is used to approximate x_d in the frequency range of this study, even if the boundary damping might be the main source. For frequencies lower than a certain frequency between f_1 and f_2, with the series in Eq. (33) truncated to the first term, C_{in} and G can be written as

$$C_{\text{in}} \approx C_0 + \frac{4\pi\epsilon_0^2 V_0^2}{j_{01}J_1(j_{01})\sigma(h-\langle x_s\rangle)^2} F_C(f) \times \int_0^a J_0\left(\frac{j_{01}}{a}r\right)\frac{r\,dr}{(h-x_s)^2}, \tag{50}$$

and

$$G \approx \frac{4\pi\epsilon_0^2 V_0^2}{j_{01}J_1(j_{01})\sigma(h-\langle x_s\rangle)^2} F_G(f) \int_0^a J_0\left(\frac{j_{01}}{a}r\right)\frac{r\,dr}{(h-x_s)^2}, \tag{51}$$

where

$$F_C(f) = \frac{f_1^2 - f^2}{4\pi^2(f_1^2-f^2)^2 + (2\alpha_1 f)^2}, \tag{52}$$

$$F_G(f) = \frac{2\alpha_1 f^2}{4\pi^2(f_1^2-f^2)^2 + (2\alpha_1 f)^2}, \tag{53}$$

and

$$f_1 = \frac{j_{01}c}{2\pi a}\sqrt{1-\frac{a^2 S_d}{j_{01}^2 T}} = \frac{j_{01}c}{2\pi a}\sqrt{1-\frac{a^2\epsilon_0 V_0^2}{j_{01}^2 T(h-\langle x_s\rangle)^3}}. \tag{54}$$

When $f=f_1$, $F_C(f)=0$, and $C_{\text{in}}=C_0$, because $\mathbf{x_{dA}}\exp(j\omega t)$ lags behind $v_0\exp(j\omega t)$ by $\pi/2$ [see Eq. (34)], i.e., no real part of $\mathbf{x_{dA}}$ is in phase with v_0. Thus $\mathbf{x_{dA}}$ makes no contribution to C_{in}. The frequencies corresponding to the maximum and minimum of $F_C(f)$ can be found by setting $dF_C(f)/df=0$:

$$f_{a,b} = f_1\sqrt{1\mp\frac{\alpha_1}{\pi f_1}} \approx f_1 \mp \frac{\alpha_1}{2\pi}. \tag{55}$$

When $f=f_a$, $\mathbf{x_{dA}}\exp(j\omega t)$ lags behind $v_0\exp(j\omega t)$ by about $\pi/4$, the real part of $\mathbf{x_{dA}}$ is positive and makes a maximum contribution to C_{in}. When $f=f_b$, $\mathbf{x_{dA}}\exp(j\omega t)$ lags behind $v_0\exp(j\omega t)$ by about $3\pi/4$, the real part of $\mathbf{x_{dA}}$ makes a maximum negative contribution to C_{in}. The closer f is to f_a or f_b, the bigger is the absolute value of $F_C(f)$.

The difference between C_{in} and C_0 is obvious. While C_0 is only a function of x_s, C_{in} is related to both x_s and x_d. In Eq. (10), although C_1 is small, its effect on C_{in} can be large, because there is a large coefficient V_0/v_0. Actually, C_1 is proportional to $V_0 v_0$, hence the term $(V_0/v_0)\text{Re}(\mathbf{C}_1)$ is proportional to V_0^2, and is independent of v_0. Therefore, regardless of how small v_0 is, the effect of C_1 on C_{in} cannot be ignored, except for: (1) $f=f_n$; (2) a certain frequency between f_n and f_{n+1}; and (3) very high frequencies where the integration of the last term of Eq. (10) results in a very small value.

It can be shown that the relation between C_{in} and V_0^2 in Eq. (50) is consistent with Beranek's result for the free capacitance of condenser microphones [by finding a solution for C_{E0} from Eqs. (6.32) and (6.33) in Ref. 2]. However, Eq. (50) contains more information, since $F_C(f)$ may be positive, zero, or negative, depending on the value of f.

III. EXPERIMENTS AND COMPARISON

A. Some physical quantities

Before calculating the theoretical values of C_{in}, the following physical quantities were found experimentally: (1) $a=50.80$ mm was measured directly; (2) $\sigma=0.1009$ kg/m^2 came from the mass of the diaphragm material divided by its area before the ring was bonded onto the material; (3) $T=1967$ N/m was obtained using the admittance method;[17] and (4) $h=195.0\,\mu$m was obtained using Eq. (2), where $C_{00}=368.2$ pF was measured using a QuadTech 1693 Precision RLC Digibridge with a 1000 Hz sinusoidal test voltage $v_0\cos\omega t$. This voltage was not the 31.6 mV excitation voltage applied across the microphone during the measurement of the input capacitance. Here[18] $v_0=1.414$ V but $V_0=0$. The electrostatic force was a function of $v_0^2\cos^2\omega t = v_0^2(1+\cos 2\omega t)/2$. It had no fundamental component but a very small second harmonic and a constant component. The capacitance measured here was approximately equal to C_{00}.

The values of $\alpha_1=\pi(f_b-f_a)$ were calculated with f_a and f_b found from the measured data of C_{in} for $P=35-50\,\mu$m Hg and $V_0=10-100$ V. The values of α_1 for $P=0$ were obtained by taking an average of α_1 for different values of P, since α_1 does not change significantly with P at such a low air pressure. Baker *et al.* have reported a similar result showing that, when $P<1$ mm Hg, essentially all air damping is removed.[14] However, α_1 varies with V_0 rather than keeps constant as predicted by Eq. (36). Therefore, it is possible that, according to Eq. (49), the variation in α_1 with respect to V_0 is due to the changes in R. This dependence is most likely related to the details of the construction of the microphone.[19] The diaphragm was made with a ring bonded onto a stretched membrane. The tension in the diaphragm is borne entirely by the conductive epoxy adhesive between the diaphragm and ring which causes high mechanical damping. Such an arrangement does not permit a clear boundary condition $x|_{r=a}=0$ but causes a negative, small, static displacement, because the shear force couple acting on the surfaces of the adhesive layer has turned the inner edge of the adhesive layer to a position inside the ring. This initial negative static displacement has been set to zero by using the value of h measured with the bridge as described in last paragraph. When V_0 is varied from 10 to 60 V, the electrostatic attraction force helps to clean up the boundary condition, thus R decreases a little and α_1 increases from 12.7 to 13.8 s^{-1}. When V_0 is increased from 60 to 100 V, the adhesive layer is subject to a force which tends to move the diaphragm away from the ring. Thus R starts to increase and α_1 graduately decreases to 2.17 s^{-1}.

B. Input capacitance

The capacitance of condenser microphones is usually measured with an admittance bridge. To investigate the variation of input capacitance with polarization voltage and test or excitation frequency, the measurement setup should be able to provide a variable high polarization voltage and a wide and continuous frequency range. The bridge available did not satisfy these requirements. Therefore, an FFT ana-

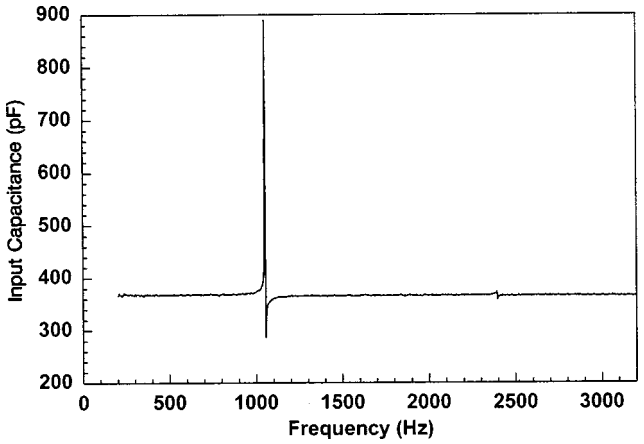


FIG. 3. Input capacitance measured as a function of frequency for $V_0 = 100\,\mathrm{V}$ and $P = 35\,\mu\mathrm{m\,Hg}$ showing the first two symmetric vibration modes of the diaphragm at 1052 Hz and 2392 Hz, respectively.

lyzer, including its built-in generator and a resistor connected to the microphone in series, were used in the setup.[17]

The values of V_0 were set from 0 to 100 V with 10 V steps. For each value of V_0, the value of P was adjusted from 35 μm Hg (the best vacuum which could be reached with the vacuum system available) to 50 μm Hg with 5 μm Hg steps to investigate the effect of the atmospheric air pressure on the measurements. The frequency range for the test signals was decreased from (0–25.6) kHz to (0–1.6) kHz and then to (1025–1075) Hz to observe the variation of the electric admittance with frequency.

The input capacitance was obtained with $C_{\mathrm{in}} = \mathrm{Im}(Y)/\omega$. In Fig. 3, the input capacitance which was measured as a function of frequency over the range of 0–3200 Hz for $V_0 = 100\,\mathrm{V}$ and $P = 35\,\mu\mathrm{m\,Hg}$ shows the first two symmetric vibration modes of the diaphragm. Each of Figs. 4(a) through (d) shows a family of experimental curves for C_{in} in the neighborhood of f_1, with V_0 varied from 0 to 100 V. The difference between the figures is related to the value of P under which the measurements were conducted. If Figs. 4(b) through (d) are compared with Fig. 4(a), the former are seen to have a similar shape but to have smaller fluctuations and a shift to higher frequencies due to more residual air loading. In these figures, each curve, except that for $V_0 = 0$, has a peak and a valley which rapidly become sharper as V_0 is increased, as predicted by Eqs. (50) and (52). Furthermore, the influence of S_d on f_1 in Eq. (54) is shown clearly in Figs. 4(a) through (d). When V_0 is increased, f_1, f_a, and f_b all shift to lower frequencies due to stronger negative stiffness.

For $P = 0$, the theoretical values of C_{in} were calculated by substituting $f = 1025 - 1075\,\mathrm{Hz}$, $V_0 = 0 - 100\,\mathrm{V}$, and the corresponding values of α_1 into Eq. (50). The data for C_{in} and f are plotted in Fig. 4(e) where a family of theoretical curves simply resemble their experimental counterparts in shape as shown in Figs. 4(a) through (d). However, Figs. 4(a) through (d) do not correspond to $P = 0$. To find all of the experimental values of C_{in} for $P = 0$ by extrapolation is tedious and unnecessary, since by observation, the expected family of curves for C_{in} found by extrapolation is not expected to change very much from those in Figs. 4(a) through (d). Thus instead of plotting the whole family of curves by extrapolation, only two frequencies, 1030 Hz and 1070 Hz, were chosen for further comparisons between the theoretical and extrapolated experimental results. These two frequencies have values close to f_1, with one lower than f_a and the other higher than f_b. The data for C_{in} were taken at 1030 Hz and 1070 Hz from Figs. 4(a) through (d) corresponding to $P = 35 - 50\,\mu\mathrm{m\,Hg}$, respectively. After extrapolation to $P = 0$, the experimental data points and the theoretical curves for C_{in} are plotted in Fig. 5 as a function of V_0. The percentage differences in the capacitance between the points and the curves are less than 1.5%.

The reason that the value of C_{in} for $P = 0$ may be found by extrapolation is as follows: When $P = 35 - 50\,\mu\mathrm{m\,Hg}$ in the microphone used for this study, the effect of the air on the diaphragm vibration mainly arises from the stiffness of the air between the diaphragm and the backplate. In addition, the tension in the diaphragm is not very large, and the velocity of the transverse waves in the diaphragm, which is about 140 m/s, is considerably less than the speed of sound in air. According to Morse[4] [Eq. (20.9)], for the frequencies close to f_1 the forced average displacement of such a microphone diaphragm, including the effect of the air inside, is inversely proportional to the air pressure, and so is the absolute value of $\mathrm{Re}(\mathbf{C}_1)$.

C. Static capacitance with bias

To avoid the diaphragm being forced into vibration during the capacitance measurements, a static method was introduced from semiconductor physics.[20,21] The measurement setup, as shown in Fig. 6, consists of a Keithley Model 230 Programmable Voltage Source, a Keithley Model 619 Electrometer, a Keithley Capacitance Standard (Part 5900-301-8) C_s connected in series with the condenser microphone under test, and a switch S for charging and discharging the microphone. The output voltage V_a of the Model 230 voltage source was very stable during the experiments and there was no need to monitor it all of the time. Therefore, the Model 619 electrometer was used to measure the voltage V_i, except that the voltage V_a was adjusted to new values for the investigation of the relationship between C_0 and $V_0 = V_a - V_i$. The value of $C_s(1820.79\,\mathrm{pF})$ was chosen to be much greater than the expected capacitance of the microphone, so that the voltage V_0 applied across the microphone was made to be as close as possible to the 100 V output limit of the Model 230 voltage source. The condenser microphone was put in the vacuum chamber to keep the measurement environment the same as that for the measurements of C_{in}. Also shown in Fig. 6 are two capacitors representing the stray capacitance, where C_w is produced by the wires across the microphone, while C_c includes both the wire capacitance and the input capacitance of the Model 619 electrometer. The purpose of this setup was to compute the capacitance of the microphone from the dc voltage distribution across the microphone under test and the capacitance standard with the stray capacitances experimentally predetermined. Since there was always a dc voltage applied across the microphone during the measurements, and there was no dynamic diaphragm motion after the free oscillation had decayed away, the quantity measured was C_0.

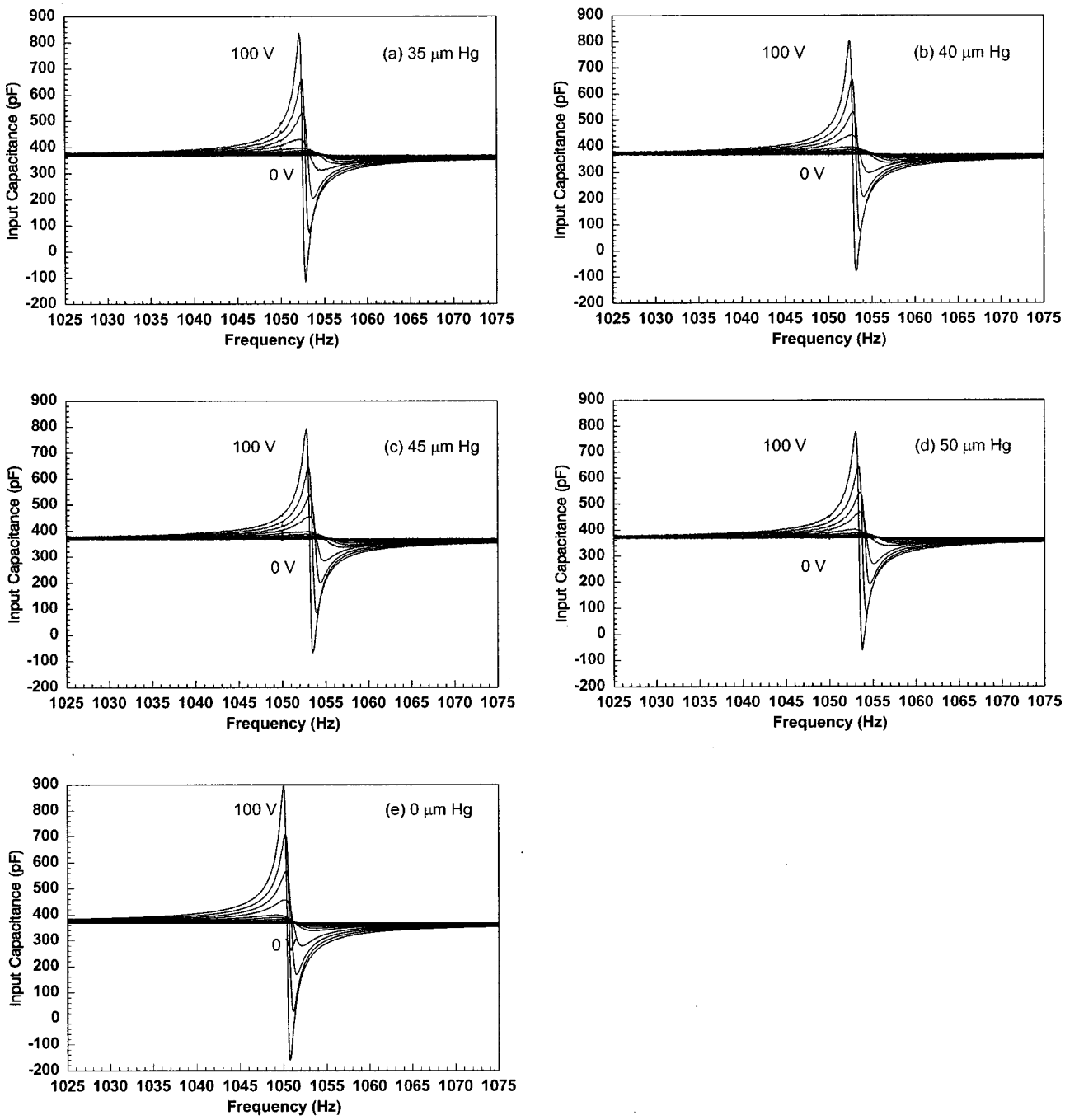


FIG. 4. Input capacitance as a function of frequency close to the first resonance of the diaphragm when the polarization voltage was varied in 10 V steps, measured with an air pressure of (a) 35, (b) 40, (c) 45, and (d) 50 μm Hg, respectively, and (e) calculated with zero air pressure.

The stray capacitances C_w and C_c were found using two groups of the Keithley Capacitance Standards, successively, to replace the microphone in which the capacitance C_0 was unknown. First, the voltages V_i and $V_0 = V_a - V_i$ were measured for the Keithley parts 5900-301-5, 46.424 pF and 5900-301-6, 180.217 pF, connected in parallel in place of the microphone. Then, the corresponding voltages were measured for the Keithley Part 5900-301-7, 463.885 pF, in place of the previous Keithley Parts. From these data, $C_w = 37.8$ pF and $C_c = 364.5$ pF were obtained.

Finally, C_0 was measured by reading V_a and V_i after the Keithley Capacitance Standard was exchanged with the microphone:

$$C_0 = \frac{V_i}{V_a - V_i}(C_s + C_c) - C_w = \frac{V_i}{V_0}(C_s + C_c) - C_w. \quad (56)$$

The experimental data points for C_0 and V_0 are plotted in Fig. 5. The theoretical curves for C_0 as a function of V_0 are also plotted there. The percentage differences between the experimental and theoretical results are less than 1%.

D. Comparison between input capacitance and static capacitance with bias

Figure 5 shows that the magnitude of the difference between C_{in} and C_0 tends to increase with V_0, and that the sign

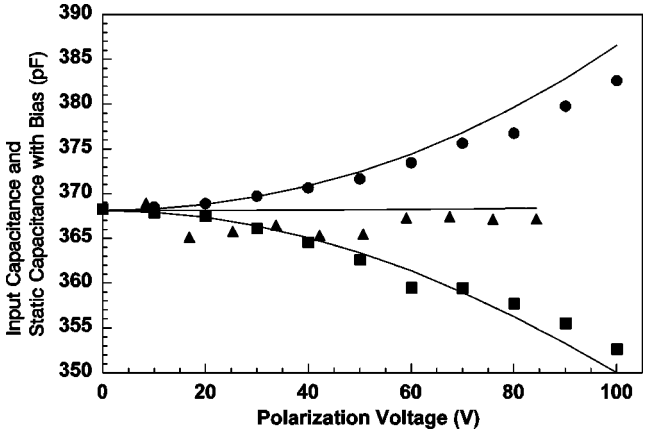


FIG. 5. Input capacitance and static capacitance with bias as a function of polarization voltage for zero air pressure. Solid lines—theoretical; data points—experimental.

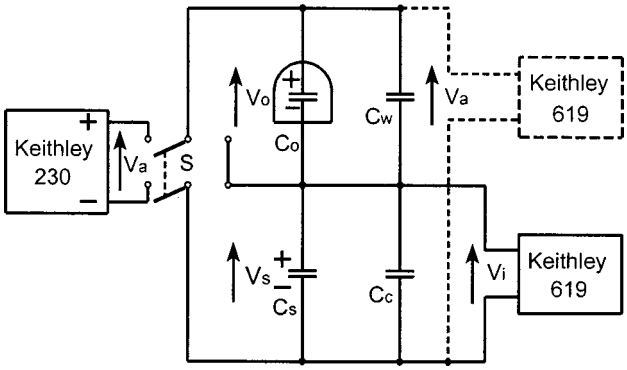


FIG. 6. Setup for measuring the static capacitance of condenser microphones varying with polarization voltage.

of the difference has two cases as predicted by Eqs. (50) and (52): (1) positive when the test frequency (i.e., 1030 Hz) is lower than f_1; and (2) negative when the test frequency (i.e., 1070 Hz) is higher than f_1. This is probably why some researchers have obtained a systematic error during their measurements of the static capacitance with bias. In the plots of capacitance as a function of the polarization voltage reported in one study,[6,7] the predicted curves are C_0 while the measured data points are actually C_{in}. The percentage differences between these results are as much as 5%. Such differences are very similar to those found in Case (1) above. The results reported in another study[8] cover both cases, as shown in Table I. For the microphones with a polyethylene diaphragm, the test frequency used was probably higher than f_1 and f_b but close in value. The percentage differences between the analytical and experimental capacitances are negative and have a magnitude as large as 11%. For the microphones with a metal diaphragm, the test frequency used was probably lower than f_1 and f_a and probably not very close in value. The percentage differences in capacitance are positive and less than 2.7%. This can be explained with Eq. (52).

IV. CONCLUSIONS

In general, the capacitance of a polarized condenser microphone measured with an ac signal is the input capacitance, since diaphragm motion exsists. An expression for the input capacitance is derived in this paper in terms of the dynamic displacement of the diaphragm. The dynamic displacement with damping can be approximated with sufficient accuracy by the first normal mode. The expression for the input capacitance conforms to Beranek's result for the free capacitance of condenser microphones, but it presents more information. Both the frequency characteristics and the polarization voltage dependence of the capacitance measured with the ac method are very close to those predicted by the theory presented.

A method for measuring the static capacitance of polarized condenser microphones is presented. This method is based on the distribution of dc voltages across an unknown capacitor and a known capacitor connected in series. Since the diaphragm of the condenser microphone under test is not excited by any ac test voltage, the quantity measured is the true static capacitance of the polarized condenser microphone. Comparisons between the experimental and theoreti-

TABLE I. Experimental values of the capacitance of some condenser microphones reported in the literature[a] and analyzed by the current authors.

Microphone number	Diaphragm material	f_1 (Hz)	Capacitance (pF) Experimental (C_{in}, actually)	Capacitance (pF) Analytical (C_0)	Percentage difference[b] (%)
111 070	Polyethylene	1325	11.5	12.64	−9.0%
92 270	Polyethylene	1797	12.4	13.67	−9.3%
6370	Polyethylene	1875	9.5	10.74	−11%
		2000 (test frequency)[c]			
11 670	Metal	4818	20.03	19.5	2.7%
42 270	Metal	7706	18.9	18.4[d]	2.7%
82 170	Metal	8320	36.4	36.1	0.8%
61 170	Metal	10 910	30.4	30.3	0.3%

[a]See Ref. 8.
[b]Calculated by the current authors.
[c]The test frequency used is not reported. It was probably between 1875 and 4818 Hz, and is assumed to have been 2000 Hz.
[d]The original wrong value, 17.6 pF, has been corrected by the current authors.

cal values for the static capacitance with bias indicate that the principle of this static method is correct.

The experimental results show that the magnitude of the difference between the input capacitance and the static capacitance with bias tends to increase with increasing polarization voltage. In addition, the magnitude and the sign of the difference depends on the relative values of the test frequency and the resonance frequency. These phenomena, including the results reported by other researchers, can be explained by the theory presented in this paper.

ACKNOWLEDGMENTS

The generous financial support provided by Dr. Jing-Yau Chung is gratefully acknowledged. The authors would also like to thank Dr. Yonhua Tzeng, Dr. Chin-Che Tin, Dr. Ralph Hing-Chung Zee, and Donald Lester Sirois for assistance with the use of experimental equipment.

[1] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics*, 3rd ed. (Wiley, New York, 1982).
[2] L. L. Beranek, *Acoustics* (American Institute of Physics, New York, 1986), paperback ed.
[3] F. V. Hunt, *Electroacoustics* (American Institute of Physics, New York, 1982), paperback version.
[4] P. M. Morse, *Vibration and Sound* (American Institute of Physics, New York, 1986), paperback ed.
[5] American National Standard S1.1-1994, *Acoustical Terminology* (United States of America Standards Institute, New York, 1994).
[6] I. J. Busch-Vishniac, ''Response of an edge-supported circular membrane electret earphone. Part II—Experimental results,'' J. Acoust. Soc. Am. **75**, 990–995 (1984).
[7] I. J. Busch-Vishniac, ''Response of an edge-supported, annular, foil electret earphone,'' J. Acoust. Soc. Am. **81**, 1179–1189 (1987).
[8] J. E. Warren, ''Contributions to the mathematical modeling of capacitance microphones,'' Ph. D. thesis, Purdue University, Lafayette, IN (1971).
[9] L. L. Beranek, *Acoustical Measurements* (American Institute of Physics, New York, 1988), revised ed.
[10] American National Standard S1.10-1966 (R 1986), *Method for the Calibration of Microphones* (United States of America Standards Institute, New York, 1986), revision.
[11] Brüel & Kjær, *Condenser Microphones and Microphone Preamplifiers for Acoustic Measurements, Data Handbook* (Naerum, Denmark, 1982), revision. In addition, for example, Lawson gives the value of their condenser microphone capacitance at 250 Hz in the calibration chart.
[12] I. J. Busch-Vishniac, ''Response of an edge-supported circular membrane electret earphone. Part I—Theory,'' J. Acoust. Soc. Am. **75**, 977–989 (1984).
[13] M. S. Hawley, F. F. Romanow, and J. E. Warren, in *AIP Handbook of Condenser Microphones*, edited by G. S. K. Wong and T. F. W. Embleton (American Institute of Physics, New York, 1995), Chap. 2, pp. 8–34, Eq. (5.8).
[14] W. E. Baker, W. E. Woolam, and D. Young, ''Air and internal damping of thin cantilever beams,'' Int. J. Mech. Sci. **9**, 743–766 (1967).
[15] S. Timoshenko, D. H. Young, and W. Weaver, Jr., *Vibration Problems in Engineering*, 4th ed. (Wiley, New York, 1974).
[16] G. Chen and J. Zhou, *Vibration and Damping in Distributed Systems* (CRC Press, Boca Raton, 1993), Vol. II.
[17] X. Yan and M. J. Crocker, ''A method for measuring diaphragm tension of condenser microphones using electric admittance,'' J. Acoust. Soc. Am. **108**, 2145–2150 (2000).
[18] QuadTech, *1693 Precision RLC Digibridge Instruction Manual* (Bolton, 1992).
[19] This suggestion was made by one of the reviewers of this paper and it has been accepted by the authors.
[20] K. Ziegler and E. Klausmann, ''Static technique for precise measurements of surface potential and interface state density in MOS structures,'' Appl. Phys. Lett. **26**, 400–402 (1975).
[21] E. H. Nicollian and J. R. Brews, *MOS (Metal Oxide Semiconductor) Physics and Technology* (Wiley, New York, 1982).

A method for measuring the diaphragm tension of condenser microphones using electric admittance

Xinche Yan[a] and Malcolm J. Crocker
Department of Mechanical Engineering, Auburn University, Alabama 36849

(Received 3 January 2000; revised 20 May 2000; accepted 29 July 2000)

This paper presents a method to measure the diaphragm tension of a condenser microphone where the fundamental frequency of the diaphragm is found from the input conductance of the microphone as a function of frequency. In this contact-free, purely electrical approach only the diaphragm and the backplate of the microphone are used as the main system for exciting and detecting the diaphragm vibration. © *2000 Acoustical Society of America.* [S0001-4966(00)00711-6]

PACS numbers: 43.38.Bs, 43.40.At, 43.40.Dx [SLE]

LIST OF SYMBOLS

a	radius of diaphragm and backplate (m)
C_A, C_B, C_I	values of capacitors (F)
$F_G(f)$	frequency factor (s)
f	frequency (Hz)
f_1, f_2	natural frequencies for symmetric modes of diaphragm vibration (Hz)
$f_1^{(0)}$	fundamental frequency for symmetric free vibration of diaphragm without bias, damping, and air loading (Hz)
G	input conductance ($S=\Omega^{-1}$)
G_{AA}, G_{BB}	autospectra (V^2/Hz)
G_{AB}	cross-spectrum (V^2/Hz)
h	static distance between diaphragm and backplate with axis of condenser microphone aligned horizontally without any voltages applied (m)
i, i_C, i_R	alternating currents (A)
$J_0(x), J_1(x), J_2(x)$	Bessel functions of order zero, one, and two
j	$(-1)^{1/2}$
j_{01}, j_{02}	roots of equation J0(x)=0 ($j01=2.4048$, $j02=5.5201$)
k_s	wave number (m^{-1})
P	air pressure in vacuum chamber (mm Hg=Torr)
R_A, R_B, R_I, R_P, R_S	values of resistors (Ω)
r	radial coordinate of diaphragm (m)
T	tension in diaphragm (N/m)
t	time (s)
$\mathbf{u}_A$	complex amplitude of diaphragm velocity (m/s)
V_0	polarization voltage (V)
v_0	amplitude of sinusoidal excitation voltage across condenser microphone (V)
v, v_A, v_B, v_C, v_G	alternating or random voltages (V)
$v_{A\mathrm{rms}}, v_{B\mathrm{rms}}$	root-mean-square values of voltage (V)
x	displacement of diaphragm (m)
x_s	static displacement of diaphragm (m)
Y	input electric admittance ($S=\Omega^{-1}$)
α_1	attenuation coefficient (s^{-1})
ε_0	permittivity of free space (8.854×10^{-12} F/m)
π	ratio of circumference to diameter of circle
σ	surface mass density of diaphragm (kg/m^2)
ϕ	phase angle (rad)
ω	angular frequency (rad/s)

INTRODUCTION

In the study of the performance of condenser microphones, sometimes the tension in the diaphragm needs to be determined. The purpose of this paper is to present a simple, contact-free, and purely electrical method for measuring the tension in the diaphragm of a condenser microphone. The methods reported in the literature[1–6] may be classified in two main categories: One is based on the relationship between the diaphragm tension and the static displacement of the diaphragm,[1–4] while the other is based on finding the tension from the resonance frequency of the diaphragm.[5,6] Some difficulties remain in the use of these methods as described below.

To determine the diaphragm tension, Wente[1] measured the deflections produced at various points on the diaphragm by a known polarization voltage, and he also measured the gap length between the diaphragm and the backplate by applying a mechanical force until the diaphragm touched the backplate. He used a device to magnify these deflections 30 000 times. This device was similar to that used in Miller's phonodeik.[7] The latter is an instrument for photographing and demonstrating sound waves by a system that is in contact with the diaphragm. Having learned this instrument in detail, one must realize how delicate the system needs to be

[a]Present address: Knowles Electronics, LLC, 1151 Maplewood Drive, Itasca, IL 60143.

for contact measurements of diaphragm deflections.

Busch-Vishniac[2] determined the diaphragm tension by finding the dc potential which caused the diaphragm of an electrostatic transducer to collapse onto the backplate with the gap length fixed by the circular ridge of a known height around the backplate. In her experiment, the diaphragms were made of FEP (fluoroethylenepropylene) or PTFE (polytetrafluoroethylene) with a thin aluminum coating on the outward side of the diaphragm. Since FEP and PTFE are good insulating materials, they may prevent a strong electric current from flowing through the electrodes of the transducer under test even during the collapse. However, if both the diaphragm and the backplate are made of metal without an insulating coating, the ''collapse'' method may cause damage to the experimental system. At the same time, if the air gap is not very small, then the dc voltage needed must be very high thus requiring a special power supply. Furthermore, for an existing transducer where the gap length is not known, it is very difficult to determine its length. For example, Warren[5] used a laser interferometer to determine the air gap by applying and then slowly releasing a vacuum to the back chamber to force the diaphragm onto the backplate and then back to the undeflected position. In addition to the complication of this optical system itself, an excellent vibration isolator was needed in order to mount this sensitive setup.

Hu *et al.*[3] measured the diaphragm tension based on the relationship between the average displacement of a clamped circular membrane and the tension in the diaphragm. They assembled a Mylar sheet as the diaphragm of a microphone. With the diaphragm clamped between an outer ring and an inner ring, a tension ring was screwed through the inner ring and forced against the diaphragm, thus stretching it. The diaphragm displacement was assumed to be equal to the axial displacement of the tension ring where the number of threads per unit length was known. In addition, they measured the stiffness of a Mylar strip in advance from an experiment similar to that to determine the Young's modulus with the use of an optical detection system. The diaphragm tension was found by relating the diaphragm displacement to the stiffness data.

The method used by Anderson *et al.*[4] is somewhat similar to that used by Hu *et al.* They screwed a threaded bushing to push the backplate against the diaphragm fixed between a retaining ring and a tension collar. At the same time, a tensile force was applied to the diaphragm via several load arms and was measured with strain gages mounted on the load arms. However, both methods will fail for those microphones in which the diaphragm tension cannot be adjusted.

In the second category of tension measurements, both Warren[5] and Zuckerwar[6] measured the resonance frequency of the diaphragm in a vacuum chamber which was used to eliminate the influence on the resonance frequency from the air loading on the diaphragm. Since it is impossible to excite or detect the diaphragm vibration acoustically in the vacuum chamber, it is necessary to develop an alternate method to solve this problem.

To excite the diaphragm into vibration, Warren[5] applied a sinusoidal voltage to his microphone cartridge. Busch-Vishniac[4] also applied an electric signal to her electrostatic earphone. Their excitation systems were simple, because, beside the internal construction of the transducer, i.e., the diaphragm and the backplate, no additional device was used. To detect the diaphragm vibration, Warren[5] used a fiber-optic sensor. Zuckerwar[6] detected the diaphragm resonance frequency in a vacuum chamber with a microphone carrier system. Might there be any other approaches that could be used to detect the diaphragm vibration?

Beside the above methods, electromagnetic excitation and detection as well as capacitive detection might be other choices. For example, the Brüel & Kjær Magnetic Transducer Type MM0002 can be used to excite or detect the vibration of an object as a noncontacting device.[8] Unfortunately, the electromagnetic excitation only works for ferromagnetic objects, although the magnetic transducer MM0002 is also sensitive to the motion of nonferromagnetic conducting materials such as aluminum and copper, due to the formation of eddy currents.[8] For nonferromagnetic and nonconducting diaphragms, such as those made of polymers, some high permeability disks should be glued onto the diaphragms. In this case, however, the technique is no longer contact-free. On the other hand, in addition to its use for capacitive excitation with an electrostatic actuator,[6] the Brüel & Kjær Capacitive Transducer Type MM0004 can be used for contact-free displacement detection.[8] Like electromagnetic methods, capacitive methods are not ideal either, because they increase the amount of equipment needed for the experimental setup. Since the microphone itself is a capacitance transducer, why is an additional capacitance transducer needed?

From the previous discussion about tension measurements, one may try to find an ideal method for measuring diaphragm tension which is (1) noncontacting, (2) purely electrical rather than acoustical, mechanical or optical, and (3) simple, i.e., one that makes use of the internal construction of a condenser microphone for both diaphragm excitation and detection purposes.

I. THEORY

It is encouraging that there is indeed a solution to the problem discussed in last paragraph, that is, the measurement of the electric admittance of the condenser microphone cartridge. During such a measurement the diaphragm is excited by an alternating input voltage $v_0 \cos \omega t$, and the resonance phenomena of the diaphragm become visible in the output current due to the electromechanical coupling. As shown in a companion paper,[9] the input conductance of the condenser microphone is a function of the diaphragm velocity:

$$G = \frac{V_0}{v_0} 2\pi \epsilon_0 \int_0^a \frac{\mathrm{Re}(\mathbf{u}_A) r\, dr}{(h - x_s)^2}. \tag{1}$$

Thus the resonance frequencies can be found from the peaks of the curve for the input conductance which is a function of frequency. This method is not only a noncontacting approach but it is also a simple electrical technique.

The microphone used in the work reported in this paper is the same as that used in Ref. 9: The diameters of the

diaphragm and the backplate of the microphone are the same, the backplate is solid and flat, and the microphone is placed in a vacuum chamber with the axis of the microphone aligned horizontally. It is assumed that the diaphragm is damped and its boundary is rigid. For the frequencies below a value higher than f_1 but still far lower than f_2, the input conductance of such a condenser microphone without air loading is given by[9]

$$G=\frac{4\pi\epsilon_0^2V_0^2}{j_{01}J_1(j_{01})\sigma(h-\langle x_s\rangle)^2}F_G(f)\int_0^a J_0\left(\frac{j_{01}}{a}r\right)\frac{r\,dr}{(h-x_s)^2}, \tag{2}$$

where

$$x_s\approx\frac{h}{2}\left(\frac{J_0(k_sr)}{J_0(k_sa)}-1\right), \tag{3}$$

$$\langle x_s\rangle\approx\frac{h}{2}\frac{J_2(k_sa)}{J_0(k_sa)}, \tag{4}$$

$$k_s=\sqrt{\frac{\epsilon_0V_0^2}{Th^3}}, \tag{5}$$

$$F_G(f)=\frac{2\alpha_1f^2}{4\pi^2(f_1^2-f^2)^2+(2\alpha_1f)^2}, \tag{6}$$

$$f_1=f_1^{(0)}\sqrt{1-\frac{a^2\epsilon_0V_0^2}{J_{01}^2T(h-\langle x_s\rangle)^3}}, \tag{7}$$

and

$$f_1^{(0)}=\frac{J_{01}}{2\pi a}\sqrt{\frac{T}{\sigma}}. \tag{8}$$

The quantity needed for the calculation of the diaphragm tension is the fundamental frequency of the symmetric free vibration of the undamped diaphragm without bias, as shown in Eq. (8). However, the fundamental frequency f_1 obtained from the input conductance is a function of V_0. Since f_1 approaches $f_1^{(0)}$ when V_0 approaches zero, it is possible to extrapolate f_1 to $f_1^{(0)}$ at $V_0=0$. If V_0 is not very high, hence x_s is much less than h, by plotting f_1^2 against V_0^2 as a straight line, $(f_1^{(0)})^2$ can be determined from the y-intercept of this line.

$$f_1^2=(f_1^{(0)})^2-\frac{\epsilon_0}{4\pi^2\sigma(h-\langle x_s\rangle)^3}V_0^2. \tag{9}$$

Then the diaphragm tension is given by

$$T=\left(\frac{2\pi af_1^{(0)}}{j_{01}}\right)^2\sigma. \tag{10}$$

II. EXPERIMENTS

A. Measurement setup

The capacitance of condenser microphones is usually measured with an admittance bridge. To find the fundamental frequency of the diaphragm, the admittance measurement setup should be able to provide a wide and continuous frequency range and a variable but high polarization voltage. If the test frequency is not continuous, then the peaks and valleys in the measured curve for the admittance as a function of frequency will not show up clearly. If the polarization voltage is too low, then the assumption that the excitation voltage is much less than the polarization voltage will not be correct, and the relationship between the resonance frequency and the polarization voltage will not be accurate. The bridge available did not satisfy these measurement requirements. However, these problems were solved by the use of an alternative setup with an FFT analyzer, as shown in Fig. 1.

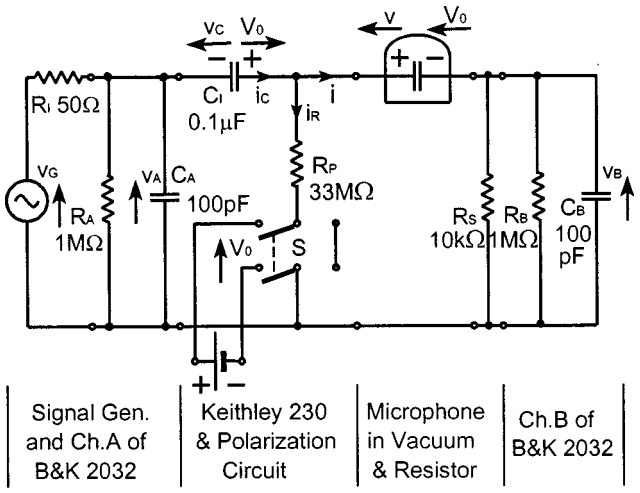


FIG. 1. Setup for measuring the input electric admittance of condenser microphones.

The measurement setup was used to find the ratio of the alternating current flowing through the microphone, i, to the alternating voltage applied across it, v. The setup consisted of a Brüel & Kjær Dual Channel Signal Analyzer Type 2032, a Keithley Model 230 Programmable Voltage Source, a polarization circuit, and a resistor R_S for current measurement which was connected in series with the condenser microphone cartridge placed in a vacuum chamber. In the polarization circuit, R_P was a protective resistor for isolating the polarization source from the alternating current; while C_I was a blocking capacitor for preventing the direct current from flowing through the Signal Generator and the Channel A of the Analyzer. A switch S was used for charging and discharging the microphone. The parallel combinations, $R_A\|C_A$ and $R_B\|C_B$, represent the input impedance of the Channel A and Channel B of the Analyzer, respectively, and R_1 the output impedance (internal resistance) of the Signal Generator.

B. Data acquisition

The polarization voltage was set by the Model 230 Voltage Source from 0 to 100 V with 10 V steps. For each value of the polarization voltage, the air pressure in the vacuum chamber was adjusted from 35 μm Hg (the best vacuum that could be reached for the vacuum system available) to 50 μm Hg with 5 μm Hg steps to investigate the effect of the air pressure on the measurements.

The test signal v_A applied to the polarization circuit including the microphone was obtained from a pseudo-random noise v_G which was the voltage of the Generator. While the

voltage v_A was measured with Channel A, the voltage v_B across the resistor R_S was measured with Channel B for determining the current i flowing through the microphone cartridge caused by the voltage v_A. The time records of the voltages v_A and v_B became two autospectra, G_{AA}, G_{BB}, and a cross-spectrum, G_{AB}, after the fast Fourier transform had been completed. Each line in the autospectra represented a sinusoidal voltage at a specific frequency. The root-mean-square values of the voltages v_A and v_B, i.e., $v_{A\text{rms}}$ and $v_{B\text{rms}}$ were obtained from G_{AA} and G_{BB}, respectively, and the phase angle ϕ between v_A and v_B at each frequency was given by the phase of G_{AB}. These data were acquired directly from the memory of the analyzer and saved on floppy disks. The interval between the two nearest spectral lines, i.e., the frequency resolution of the Analyzer, was 1/800 of its frequency band.[10] For example, there was a 4 Hz frequency interval for the frequency range 0–3200 Hz, and a 0.0625 Hz interval for the 50 Hz zoom measurements.

The magnitude of the frequency response for the voltage ratio v_B/v_A was shown on the screen of the Analyzer. This ratio is a quantity close to that of the admittance, because v_B is approximately equal to iR_S, and v_A is approximately equal to the alternating voltage v across the microphone. Thus v_B/v_A is approximately proportional to the admittance. Therefore, the frequency response of v_B/v_A can provide some qualitative information about the frequency characteristic of the admittance. When the frequency band was decreased from (0–25.6) kHz to (0–1.6) kHz, this procedure was used to estimate the resonance frequencies of the diaphragm. In this step, after the polarization voltage was applied and the air pressure in the vacuum chamber was decreased, the resonance phenomena in the frequency response of v_B/v_A were observed on the screen. The first and second resonance frequencies obtained were 1052 Hz and 2392 Hz. Although the second resonance was relatively weak, it appeared steadily during each measurement. The ratio of these two frequencies was 2.274. Comparing this ratio with $j_{02}/j_{01}=2.295$, and considering the fact that the above measurement will only give a rough estimate of the admittance, and that the frequency resolution is 4 Hz for the frequency band of 0–3200 Hz, one can have confidence that these two resonance frequencies represent the first two radial vibration modes of the diaphragm. Finally, a zoom measurement from 1025 to 1075 Hz was made around the fundamental frequency f_1 to locate the fundamental frequency more accurately.

C. Data analysis

The input electric admittance of the microphone is given by $Y=i/v$, where $v=v_A-v_B-v_C$, and

$$i=v_B\left(\frac{1}{R_S}+\frac{1}{R_B}+j\omega C_B\right). \tag{11}$$

Assuming that the instantaneous voltage v_A has a zero initial phase angle, then, for each frequency, v_A and v_B can be written, respectively, as

$$v_A=\sqrt{2}v_{A\text{rms}}e^{j\omega t}, \tag{12}$$

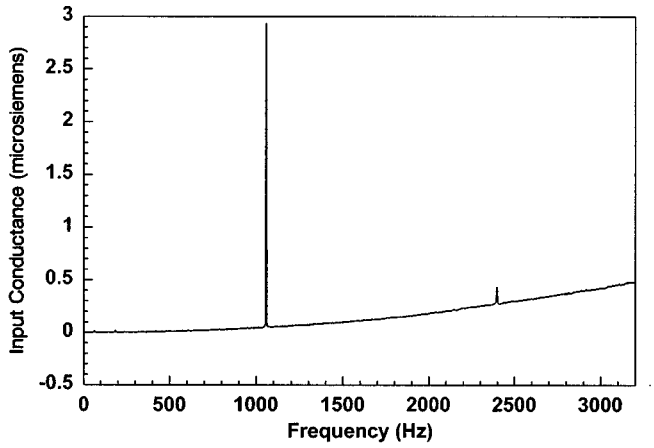


FIG. 2. Input conductance as a function of frequency measured for a polarization voltage of 100 V and an air pressure of 35 μm Hg showing the first two symmetric modes of the diaphragm at 1052 Hz and 2392 Hz, respectively.

and

$$v_B=\sqrt{2}v_{B\text{rms}}e^{j(\omega t+\phi)}, \tag{13}$$

while

$$v_C=\frac{v_A+iR_P}{1+j\omega C_I R_P}. \tag{14}$$

With the following quantities measured, $C_B=100$ pF, $C_I=0.102\,\mu$F, $R_B=1$ MΩ, $R_S=10.1$ kΩ, and $R_P=33.0$ MΩ, the electric admittance was obtained from:

$$Y=\frac{i(1+j\omega C_I R_P)}{(v_A-v_B)(1+j\omega C_I R_P)-(v_A+iR_P)}. \tag{15}$$

The input conductance $G=\text{Re}(Y)$ was calculated from the data acquired for the frequency of each spectral line and each value of the polarization voltage and air pressure.

Figure 2 shows the input conductance over a frequency band of 0–3200 Hz with a polarization voltage of 100 V and an air pressure of 35 μm Hg. There are two peaks at 1052 Hz and 2392 Hz, respectively, which represent the first two symmetric vibration modes of the diaphragm. Figures 3(a) through (d) were obtained from the zoom measurements. Each of them shows a family of curves for the input conductance as a function of frequency in the neighborhood of the fundamental frequency when the polarization voltage is varied from 0 to 100 V. As predicted with Eq. (2), each curve, except that for $V_0=0$, has a peak which rapidly becomes sharper as V_0 is increased. The peak frequency, i.e., f_1, was located from each of these curves. The influence of V_0 on f_1 is shown clearly in each figure as predicted with Eq. (7). When V_0 is increased, f_1 shifts to lower frequencies. On the other hand, each figure in Figs. 3(a) through (d) corresponds to a different value of the air pressure between 35 and 50 μm Hg, respectively. These figures have a similar shape but as the air pressure is increased the height of the peaks decreases and the peaks shift to higher frequencies due to more residual air loading. Here the residual air loading is mainly

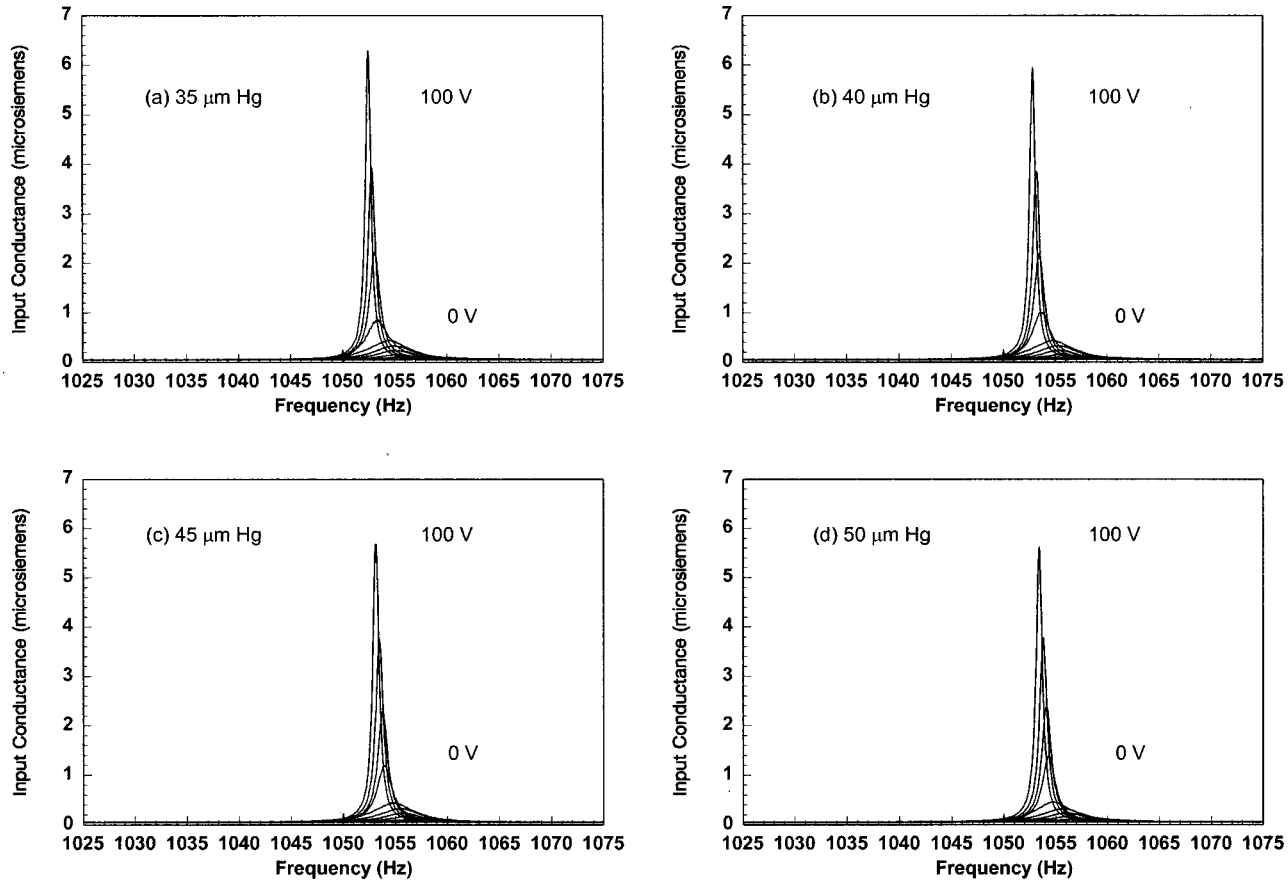


FIG. 3. Input conductance as a function of frequency close to the first resonance of the diaphragm with the polarization voltage varied in 10 V steps, measured with an air pressure of (a) 35 μm Hg, (b) 40 μm Hg, (c) 45 μm Hg, and (d) 50 μm Hg, respectively.

caused by the stiffness of the air between the diaphragm and the backplate which is much less than the stiffness of the diaphragm itself.

Morse[11] has pointed out that when the velocity of transverse waves in the diaphragm (in the case of this study, 140 m/s) is considerably smaller than the speed of sound in air and the stiffness of the air confined in the microphone is small compared to the diaphragm tension, which acts as a restoring force on the diaphragm; then the allowed frequencies for the diaphragm plus the enclosed air are linearly related to the air density [Eq. (19.18)]. Therefore, for each value of the polarization voltage, the fundamental frequency f_1 corresponding to zero air pressure can be found by extrapolation. Figure 4 shows the extrapolation procedure for the fundamental frequency when the polarization voltage is varied from 70 V to 100 V. The data points and regression lines corresponding to the other values of the polarization voltage (not shown) have a similar slope but a higher y-intercept.

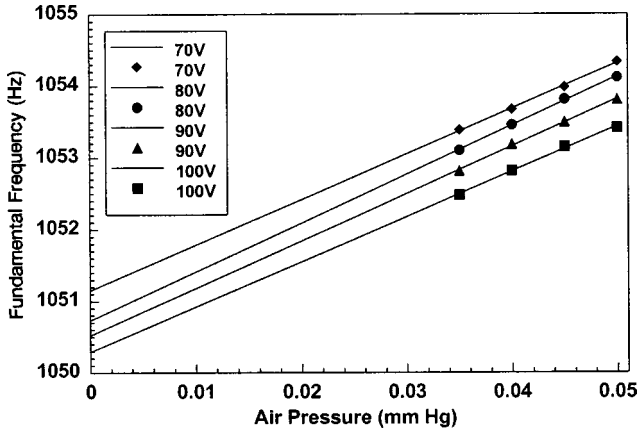


FIG. 4. Experimental data points and corresponding regression lines for the fundamental frequency as a function of air pressure, when the polarization voltage is varied from 70 V to 100 V, showing the extrapolation of the fundamental frequency to zero air pressure.

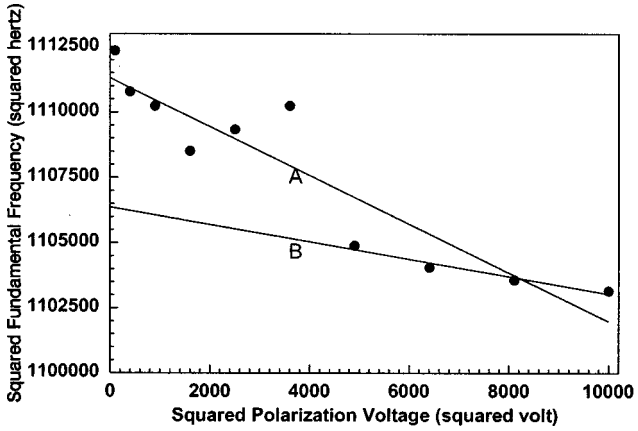


FIG. 5. Experimental data points and regression lines for the squared fundamental frequency as a function of the squared polarization voltage for a zero air pressure showing the y-intercept of Line B.

Figure 5 is plotted with the squared polarization voltage and the corresponding squared fundamental frequency for a zero air pressure obtained from the above extrapolation procedure. In Fig. 5, Line A is a regression line for all the data points. The value of $(f_1^{(0)})^2$ is obtained from the *y*-intercept of Line A, as described in Sec. I. The fundamental frequency $f_1^{(0)}$ obtained was 1054.2 Hz. To verify if this value is correct, the slope of Line A can be checked. According to Eq. (9), the theoretical slope is $-0.30\ \mathrm{Hz^2\,V^{-2}}$. However, the experimental slope is $-0.93\ \mathrm{Hz^2\,V^{-2}}$. This discrepancy is due to the poor data for the polarization voltage below 60 V. The possible reason is that the diaphragm bonded on the ring with epoxy adhesive does not permit a clear boundary condition for the polarization voltage lower than 60 V, as explained in Ref. 9. Therefore, Line B was drawn using only the data points for the polarization voltage ranging from 70 to 100 V. The slope of Line B is $-0.34\ \mathrm{Hz^2\,V^{-2}}$, which is close to the predicted value, $-0.30\ \mathrm{Hz^2\,V^{-2}}$. The fundamental frequency $f_1^{(0)}$ obtained from the *y*-intercept of Line B was 1051.9 Hz. This value, 1051.9 Hz, was considered to be more accurate and was used to calculate the tension. Finally, by substituting the values of $a=50.80\times10^{-3}$ m and $\sigma=0.1009\ \mathrm{kg/m^2}$ into Eq. (10), the tension in the diaphragm was found to be $T=1967$ N/m.

The main uncertainty in measuring the diaphragm tension is due to the abovementioned extrapolation of the fundamental frequency to a zero air pressure and to a zero polarization voltage. This uncertainty can be estimated using error analysis theory.[12] In Fig. 5, the uncertainty in the *y*-intercept of Line B, $(f_1^{(0)})^2$, is 1095 $\mathrm{Hz^2}$, which is about two divisions on the *y*-axis. By substituting this value into Eq. (10), the uncertainty in the tension measurement is found to be 2 N/m. Therefore, $T=1967\pm2$ N/m. The uncertainty in $f_1^{(0)}$ can also be found to be $f_1^{(0)}=1051.9\pm0.5$ Hz.

III. CONCLUSIONS

The main ideas used in the method for determining the diaphragm tension with electric admittance are as follows: (1) During the admittance measurement the diaphragm is excited by an ac test signal; (2) the variation of the diaphragm velocity with frequency is displayed by the measured input conductance; thus (3) the fundamental frequency of the diaphragm can be found from the variation of the input conductance with frequency (see Fig. 3); then (4) the fundamental frequency is used to determine the diaphragm tension. Hence, the two resonance phenomena that are observed, which represent the first two radial modes of the vibration of the diaphragm, are a convincing proof of the theory presented in this paper. The negative slope of the straight line fitting the experimental data points of the squared fundamental frequency and squared polarization voltage also supports the theory (see Line B in Fig. 5).

This method is a contact-free and purely electrical technique. Since the measurement is only made at the electrical terminals of the condenser microphone, this approach is much simpler than other existing mechanical or optical methods for determining the diaphragm tension.

In this study, instead of using an admittance bridge, the electric admittance of the condenser microphone was measured with an FFT analyzer by connecting a resistor to the microphone in series. The purpose of using such a setup is to provide a wide and continuous frequency range for the admittance measurements and to supply the microphone under test with a variable high polarization voltage. The experiment shows that this substitution of the FFT analyzer for the admittance bridge is feasible.

ACKNOWLEDGMENTS

The generous financial support provided by Dr. Jing-Yau Chung is gratefully acknowledged. The authors would like to thank Dr. Yonhua Tzeng and Dr. Ralph Hing-Chung Zee for assistance with the use of experimental equipment.

[1] E. C. Wente, ''A condenser transmitter as a uniformly sensitive instrument for the absolute measurement of sound intensity,'' Phys. Rev. **10**, 39–63 (1917).
[2] I. J. Busch-Vishniac, ''Response of an edge-supported circular membrane electret earphone. Part II—Experimental results,'' J. Acoust. Soc. Am. **75**, 990–995 (1984).
[3] A. Hu, F. W. Cuomo, and A. J. Zuckerwar, ''Theoretical and experimental study of a fiber optic microphone,'' J. Acoust. Soc. Am. **91**, 3049–3056 (1992).
[4] M. J. Anderson, J. A. Hill, C. M. Fortunko, N. S. Dogan, and R. D. Moore, ''Broadband electrostatic transducers: Modeling and experiments,'' J. Acoust. Soc. Am. **97**, 262–272 (1995).
[5] J. E. Warren, ''Contributions to the mathematical modeling of capacitance microphones,'' Ph. D. thesis, Purdue University, Lafayette, IN (1971).
[6] A. J. Zuckerwar, ''Theoretical response of condenser microphones,'' J. Acoust. Soc. Am. **64**, 1278–1285 (1978).
[7] D. C. Miller, *The Science of Musical Sounds*, 2nd ed. (MacMillan, New York, 1922), pp. 78–81.
[8] Brüel and Kjær, *Master Catalog 1980* (Denmark).
[9] X. Yan, M. J. Crocker, and L. J. Zeng, ''Measured capacitance of a condenser microphone as a function of diaphragm displacement,'' J. Acoust. Soc. Am. **108**, 2134–2144 (2000).
[10] Brüel & Kjær, *Dual Channel Signal Analyzer Type 2032 Instruction Manual* (Nærum, Denmark, 1983), Vol. 2.
[11] P. M. Morse, *Vibration and Sound* (American Institute of Physics, New York, 1986).
[12] J. R. Taylor, *An Introduction to Error Analysis* (University Science Books, Mill Valley, 1982).

An acoustic/thermal model for self-heating in PMN sonar projectors

Natarajan Shankar and Craig L. Hom
Org. L9-21, Building 204, Advanced Technology Center, Lockheed Martin Missiles & Space, 3251 Hanover Street, Palo Alto, California 94304

(Received 8 February 2000; revised 10 July 2000; accepted 25 July 2000)

Dielectric hysteresis and a strong material temperature dependence uniquely couple the acoustic output and temperature of a sonar projector driven by electrostrictive $Pb(Mg_{1/3}, Nb_{2/3})O_3$ (PMN). Both the source level and the source of self-heating, i.e., dielectric hysteresis, dramatically decrease as the PMN driver heats. The final temperature delineates outstanding PMN transducers from mediocre PMN transducers, so accurate acoustic performance prediction requires accurate transducer temperature prediction. This study examined this self-heating phenomenon by combining an electro-acoustics model for a PMN flextensional transducer with a thermal finite element model. The sonar model calculated the source level and heat generation rate for the PMN driver as a function of temperature. This computed source level varied 12 dB over a 75 °C temperature range solely due to the temperature dependent ceramic. The heat transfer model used the computed heat rate to predict the transducer's transient thermal response. The results clearly demonstrate that the transducer reached a steady-state equilibrium temperature, where the heat generated by the PMN driver balanced the heat dissipated. While the transducer model predicted a significant temperature rise, the corresponding acoustic output still surpassed the output of an equivalent $Pb(Zr, Ti)O_3$ (PZT) transducer by 8 dB. Good agreement with experiments made on a PMN flextensional transducer validated the model. © *2000 Acoustical Society of America.* [S0001-4966(00)00311-8]

PACS numbers: 43.38.Fx [SLE]

I. INTRODUCTION

High strain PMN ceramic drivers[1,2] represent a new enabling technology for compact, high-powered sonar transducers. This improvement reflects PMN's ten-fold increase in energy density relative to PZT-8. For example, Pilgrim *et al.*[2] have measured a 5 dB source level improvement in an early prototype flextensional PMN projector over a similar device constructed from PZT (Navy Type III). More recent PMN transducers developed by Lockheed Martin have advanced significantly from this benchmark.

While encouraged by PMN's demonstrated performance, we must understand PMN's nonlinear behavior and incorporate it into the transducer design process. Hom and Shankar[1] simulated the nonlinear prestress and field dependence of PMN in a flextensional sonar transducer. However, their analysis assumed that the transducer remained at a constant temperature during operation. Recent evidence presented in this paper indicates that nonlinear thermal heating due to dielectric loss causes large temperature rises in a PMN transducer during high voltage and high frequency operation. From a transduction standpoint, the ceramics-induced strain decreases with increasing temperature, since PMN operates in a diffuse relaxor-ferrolectric transition region near its Curie temperature. Figure 1 shows the peak strain and polarization measured for a PMN material doped with $PbTiO_3$ (PT) and $BaTiO_3$ (BT) with a ± 1 MV/m electric field at 1 Hz. [The PMN composition depicted in the figure was 0.98(0.88 PMN-0.12PT)-0.02BT and had a Curie transition temperature of 45 °C (measured with a 1 kHz weak electric field[3]).] From an acoustics standpoint, this induced strain drop corresponds to a 12 dB decrease in predicted source level from 5 °C to 80 °C for the flextensional transducer studied in this paper. Clearly, the self-heating of the ceramic diver completely dominates the acoustic performance of a PMN transducer, and the final operating temperature separates an outstanding acoustic projector (11 dB above a similar PZT projector) from a mediocre device (equivalent to a similar PZT device).

The relaxor nature of PMN makes the nonlinear thermal characteristics of a PMN transducer quite different from that of a PZT transducer. While PMN has relatively low hysteresis at high fields, it still possesses a high dielectric constant and it often operates at high frequencies and voltages. This leads to enormous reactive powers and high heat rates due to dielectric dissipation. While PMN's properties degrade with temperature, the dielectric hysteresis also drops with increasing temperature, as shown in Fig. 2. At some point, the transducer must reach thermal equilibrium; the heat generated by hysteresis balances the heat removed from the transducer by conduction, radiation, and convection to the surrounding water bath. PZT transducers also experience thermal heating due to dielectric hysteresis, when driven at extremely high voltages and frequencies. However, PZT transducers typically operate far below the material's Curie temperature, so their dielectric loss does not diminish with temperature. Thermal runaway can occur in PZT transducers under these conditions. Melting of glue and seals due to extreme heat typically limit the temperature and performance of PZT transducers. In contrast, PMN appears to have a natural shut-off point at temperatures well below the melting point of the bond joints. The temperature rises in PMN transducers only

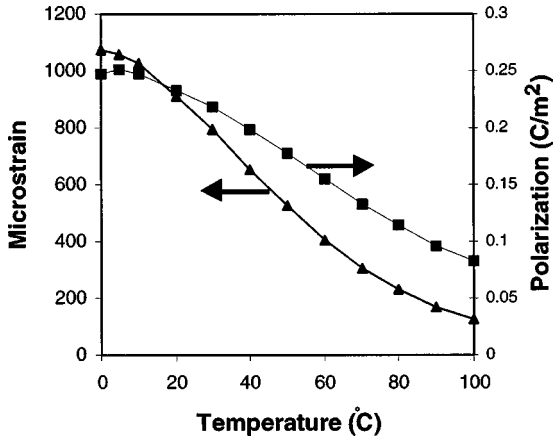


FIG. 1. Peak strain and peak polarization (±1 MV/m @ 1 Hz) for a PMN transduction material in the relaxor regime decrease sharply with temperature. Exact composition displayed is 0.98 (0.88 PMN+0.12 PT)+0.02 BT.

affect the induced strain output of the ceramic making this a problem unique to PMN.

This paper investigates the thermal heating phenomenon in a class IV flextensional PMN transducer. We combined Hom and Shankar's lumped parameter model for a PMN projector[1] with a thermal finite element model of the transducer. Accurate acoustic performance prediction required the heat transfer analysis to determine the transducer's transient temperature during operation and in turn the transducer's corresponding source level. The sonar model computed heat and source level for different temperatures. The simulations showed that significant thermal heating occurred and that the system achieved thermal equilibrium at temperatures nearly 40 °C above the ambient temperature of the surrounding water. However, the predicted acoustic output at the elevated temperature was 8 dB above that of an equivalent PZT transducer. We found that different electrical duty cycles and starting temperatures altered the transducer's acoustic performance and final temperature.

Our approach does introduce an approximation: we assumed that the ceramic had a uniform temperature in the sonar model, while the heat transfer model had a nonuniform temperature due to heat dissipation. We validated this approximation by successfully matching the computed temperature history with experiments performed on an actual flextensional transducer of the same specifications modeled.

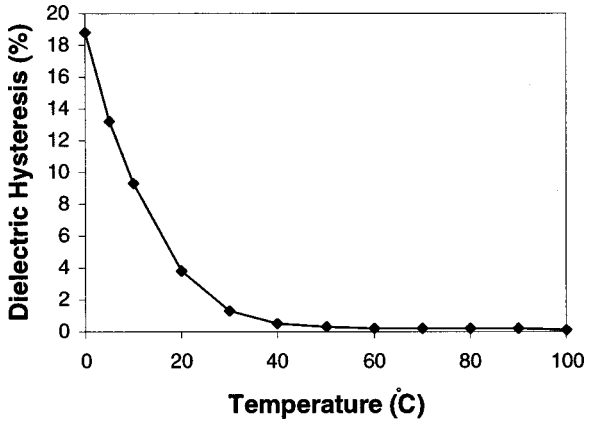


FIG. 2. Dielectric hysteresis, the source of self-heating, also declines with temperature for the PMN+PT+BT composition depicted in Fig. 1.

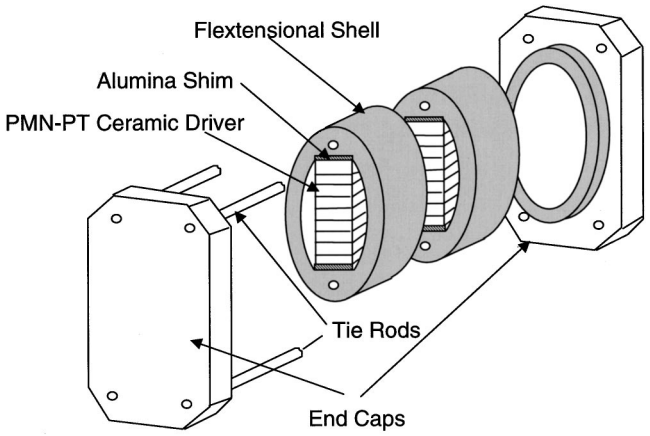


FIG. 3. A class IV flextensional sonar transducer with PMN+PT+BT stacked drivers.

II. MODEL DESCRIPTION

This investigation models the class IV flextensional transducer shown in Fig. 3. The device consists of two convex titanium shells each driven by a stack of PMN plates. The stacks have inactive alumina/silica end pieces that maintain contact with the shells. The final length of the stack was slightly larger than the shell cavity. This mismatch places the stack in residual compression (40 MPa) and keeps the stack from ''walking'' during high frequency operation. A neoprene sleeve, not shown in the figure, surrounds the shells and seals the transducer from the water.

Our approach combines a lumped-parameter class IV flextensional sonar model with a thermal finite element model as shown in Fig. 4. The sonar transducer model calculates the rate of thermal heat generated by the driver. This series of computations uses polarization, strain, and hysteresis properties measured for a single PMN composition at different temperatures. The sonar model assumes that the PMN stack had a uniform temperature. This approach gave us a series of ''snapshots'' for the heat generation rate over a broad temperature range. The thermal finite element model uses these heat generation rates as a distributed temperature dependent thermal load at each integration point. The thermal computation predicted the transient temperature distribution in the device given the duty cycle and starting temperature of the ceramic. After the finite element calculation reached thermal equilibrium, we computed the transducer's steady-state source level using the acoustic model and the average predicted steady-state temperature in the ceramic.

While the uniform temperature approximation of the sonar model significantly reduced the computational effort, it contradicts the nonuniform temperature distribution computed in the heat transfer model. Under this assumption, the temperature distribution within the ceramic has a minimal effect on the acoustic output, so a median temperature sufficiently predicts the source level and local heat rate. Clearly, we must validate this approximation, and we accomplished

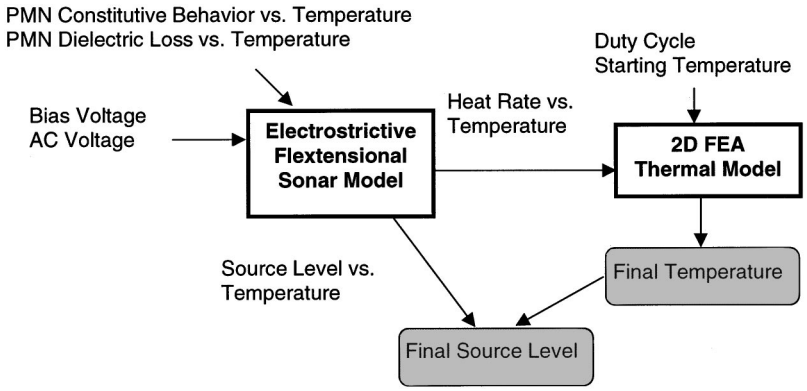


FIG. 4. Loose coupling between the acoustical model and the heat transfer model predicted the temperature history and corresponding source level for the flextensional transducer.

that by comparing our numerical results with experimental thermal and acoustic measurements made on an actual transducer under the same conditions modeled.

A. Flextensional sonar model

The electrostrictive sonar model recently published by Hom and Shankar[1] formed the basic structure of our flextensional transducer simulation. Hom and Shankar's model used a nonlinear lumped parameter model to solve for the transducer's electrical impedance and source level in the frequency domain. A one-dimensional rod vibrating along its length approximated the ceramic driver. The nonlinear, electrostrictive constitutive law for PMN developed by Hom and Shankar[4] defined the electromechanical behavior of that rod. For length extension of a PMN stack, that law reduces to the mechanical equation,

$$S(x,t)=s_{33}^{P}T(x,t)+Q_{33}P(x,t)^2, \tag{1}$$

and the electrical equation,

$$E(x,t)=-2Q_{33}P(x,t)T(x,t)+\frac{1}{2}\operatorname{arctanh}\left(\frac{P(x,t)}{P_s}\right). \tag{2}$$

x denotes the axial distance from the stack's center, and t denotes time. S, T, P, and E represent longitudinal strain, axial stress, axial polarization, and axial electric field, respectively. The material parameters Q_{33}, P_s, and k were measured experimentally from stress-free strain and polarization tests. Q_{33} is the longitudinal electrostrictive coefficient, P_s is the saturation polarization, and k is a material parameter introduced by Hom and Shankar. s_{33}^{P} represents the open-circuit elastic compliance measured from a stress–strain test.

The PMN plates were sandwiches between alternating electrodes. The voltage applied to the ''hot'' electrodes controlled the electric field in the ceramic plates. For electrostrictive ceramics, this voltage consists of a dc bias component and an ac component. The total voltage difference $\Delta\phi$ between alternating electrodes was

$$\Delta\phi(t)=V_0+\mathrm{Re}[Ve^{i\omega_0 t}], \tag{3}$$

where V_0 is the dc voltage, V is the complex ac voltage amplitude, and ω_0 is the frequency of the voltage source.

As shown in Fig. 5(a), we modified Hom and Shankar's original simulation to include the combined Helmholtz integral equation formulation (CHIEF)[5,6] for calculating the radiation impedance seen by the shell. The original model relied on an equivalent sphere approximation for the radiation load. Our new analysis uses the first five symmetric eigenmodes for the shell without the ceramic stack. A plane stress, modal finite element analysis with the commercial code ABAQUS[7] computed these eigenmodes. The first three modes were the familiar quadrupole, octopole, and membrane modes described by Butler *et al.*[8] The equivalent electrical circuit in Fig. 5(b) represents the total impedance seen by the ceramic stack. In that figure, F denotes the force applied to the stack, v^{stack} denotes the velocity of the stack, and v_j denotes the velocity contribution of the individual modes. The total impedance consists of five individual eigenmode

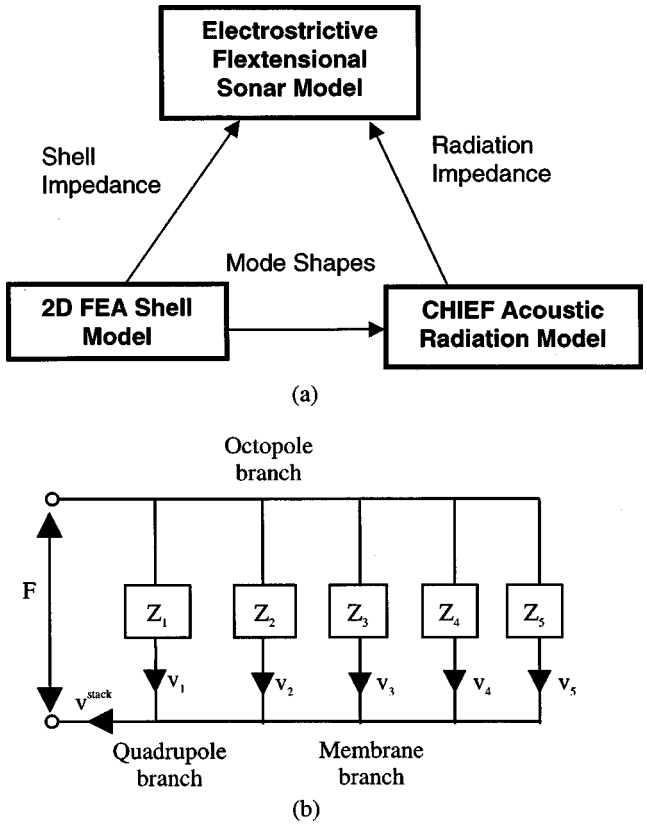


FIG. 5. (a) A combined Helmholtz integral equation formulation and modal finite element analysis was used to compute the impedance presented to the nonlinear electrostrictive stack. (b) This impedance was condensed into an equivalent circuit representing the five symmetric vibration modes of the shell and the surrounding acoustic medium.

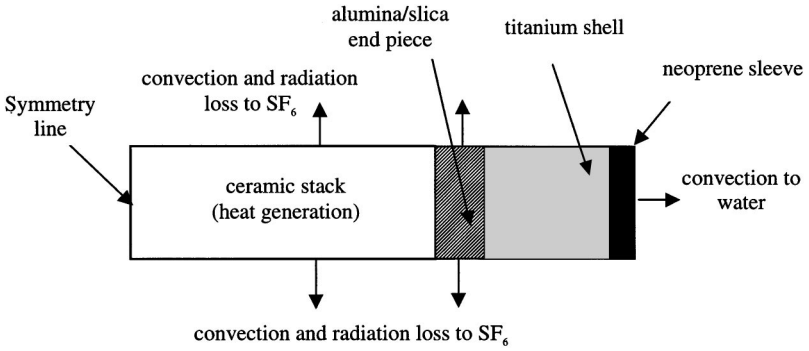


FIG. 6. The cross-sectional layout of a flextensional transducer and its heat dissipation paths.

impedances in parallel. The impedance for the jth mode equals

$$Z_j = -\frac{m_j\, Wi}{2\Gamma_j^2 \omega}\left(\Omega_j^2 - \omega^2 + \frac{i\omega\Omega_j}{Q_m}\right) + \frac{Z_j^a}{2\Gamma_j^2}. \tag{4}$$

W denotes the total length of the shell, Γ_j denotes the eigenmode displacement where the stack contacts the shell, Q_m denotes the mechanical quality factor for the shell, and ω denotes the vibration frequency. The effective mass per shell width for the jth equals

$$m_j = \int_{A^{\text{shell}}} \rho^{\text{shell}} \phi_j \cdot \phi_j \, dA, \tag{5}$$

where ρ^{shell} is the density of the shell, ϕ_j is the displacements for the jth eigenmode, and A^{shell} is the cross-sectional area of the shell.

The second term on the right hand side of Eq. (4) represents the impedance contribution seen by the stack due to acoustic radiation in water, and Z_j^a represents the acoustic radiation impedance seen by the shell. We computed Z_j^a for each eigenmode as a function of frequency using CHIEF. Then, we placed the results in Eq. (4) to compute the total impedance for a given frequency. Our process was not iterative; we assumed that the acoustic radiation did not modify the eigenmode's relative shape. However, this new approach with CHIEF should be more nearly accurate than the equivalent sphere model originally employed by Hom and Shankar.[1]

We also modified the original Hom and Shankar sonar model to include the effect of glue layers between the individual PMN plates, and the aluminia/silica end plates. The glue layers were placed in series with the ceramic stack using the method of Martin.[9] The presence of glue effectively lowered the total compliance of the ceramic stack. For the transducer modeled in this paper, the compliance of the 12.8 μm thick glue layer equaled approximately 50% of the PMN plate's compliance.

Finally, we computed the heat generation rate for the transducer directly from the polarization hysteresis. We modeled this loss as a linear electric resistor in parallel with the ceramic stack. Consequently, the electric current in the ceramic stack consists of a capacitive component and a resistive component. The sonar model computes these quantities from the predicted polarization spectrum as described by Hom and Shankar.[1] The resistive component has two contributions, one related to the acoustic load and the other related to the dielectric loss. The last part was directly proportional to the loss tangent, and we denote it as I_R. Since this component only contributes to internal heat generation, the heat rate, $\dot{Q}$, equals

$$\dot{Q} = \frac{\omega_0}{2\pi}\int_0^{2\pi/\omega_0} \Delta\phi I_R \, dt, \tag{6}$$

where $2\pi/\omega_0$ is the period of the ac voltage source. In general, we computed the internal heat generation rate for a fixed voltage frequency.

B. Thermal finite element model

Given the internal heat rate values computed by the sonar transducer model, we conducted a transient, two-dimensional, thermal analysis using the commercial finite element code ABAQUS.[7] Figure 6 shows the cross-sectional layout of the transducer and its key components. The glue bonds and internal brass electrodes were not modeled in the thermal analysis. We based this approximation on the facts that the glue layer's relative thermal resistance (thickness/conductivity) and relative heat storage capacity (thickness$\times$specific heat) equaled only 4% and 2% for a PMN plate, respectively. The glue could not store heat and conducted it quickly to the ceramic, so it was effectively thermally transparent. Inert SF_6 gas fills the shell cavity and surrounds the ceramic stack. The model consists of only half the transducer due to symmetry at the stack's center. Figure 7 shows the finite element mesh with eight-noded quadratic elements for this problem. We assumed that the thermal material behavior (i.e., specific heat and conductivity) were linear.

The internal heat was applied as a distributed heat load in the PMN elements. This situation made the boundary conditions nonlinear, since the heat load in an element depends on the local temperature. Sonar transducers do not operate continuously (a 100% duty cycle), so we simulated a pulse on/pulse off duty cycle in our analysis. For example, a 33% duty cycle could be ac voltage on for 1 s and then off for 2 s. Our transient analysis steps in time. If the transducer was pulsed on during a particular time step, the internal heat rate for an element in the PMN was computed from the sonar model results and the local temperature. That rate was applied as a distributed heat flux to the element. If the transducer was pulsed off during a particular time step, the distributed heat flux in the ceramic was zero. An ABAQUS user-

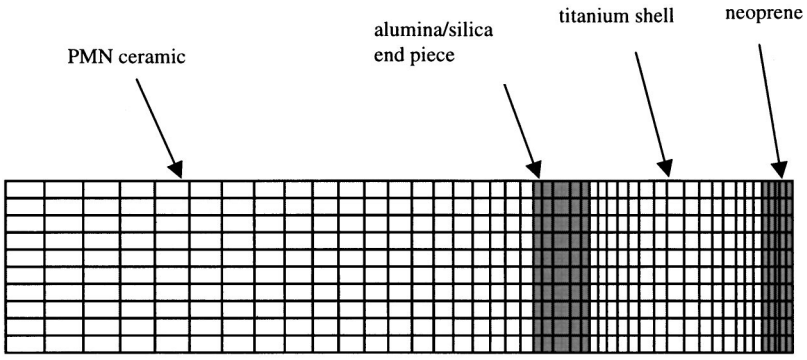


FIG. 7. Finite element mesh used for transient thermal analysis of the flextensional transducer.

subroutine was written that determined this distributed heat flux in the ceramic given the temperature and time. Using the model, we simulated the performance and temperature rise in a fully assembled unit in water, and compared these results with experimental measurements.

We assumed that heat generated in the PMN was dissipated through three mechanisms: (1) heat conducted through alumina/silica end pieces to the shell and neoprene; (2) heat radiated from the ceramic and end pieces to the inner surface of the shell; and (3) heat removed from the ceramic and end piece surfaces by induced natural convection of the SF_6 gas. The outer surface of the neoprene sleeve was in direct contact with water, so we assumed a relatively high heat transfer coefficient of 1000 W/(m^2 K) for the outer surface. At the beginning of each finite element computation, we set the temperature of the ceramic to the initial temperature measured in the experiments being simulated.

The heat flux, $\dot{q}$, for the radiation condition boundary was given by

$$\dot{q} = \varepsilon\sigma[(T - T_z)^4 - (T_{\text{ref}} - T_z)^4], \tag{7}$$

where T denotes temperature, ε denotes the emissivity of the surface, and σ denotes Boltzmann's constant. T_{ref} is a reference ambient temperature, and T_z is the value of absolute zero on the temperature scale used. We assumed an emissivity for the ceramic and the end pieces of 0.7, and a reference ambient temperature equal to the surrounding water temperature.

The heat flux for the natural convection boundary condition was given by

$$\dot{q} = h(T - T_{\text{ref}}). \tag{8}$$

The heat transfer coefficient, h, for natural convection in the SF_6 was estimated from empirical correlation with experiments conducted on thin, heated vertical plates suspended in air. Based on those experiments, we used

$$\text{Nu} = \frac{hL_r}{\lambda} = 0.59\,(\text{Gr}\,\text{Pr})^{1/4}, \quad 10^4 < \text{Gr} < 10^9. \tag{9}$$

This equation approximates the Nusselt number, Nu, as a function of the Grashoff number, Gr, and Prandtl number, Pr. L_r represents a characteristic length, which we set to the width of the ceramic stack. λ represents the thermal conductivity of the fluid. Combined with the SF_6 properties and the transducer test conditions, we used Eq. (9) to derive the heat transfer coefficient for the ceramic and alumina/silica surfaces.

III. TRANSDUCER RESULTS

We used the model to simulate a class IV flextensional, sonar transducer built by the Advanced Technology Center of Lockheed Martin Missiles & Space. Each ceramic stack in the transducer consisted of 20 PMN-PT-BT plates made from the 0.98(0.88PMN-0.12PT)-0.02BT composition depicted in Figs. 1 and 2. The plates were 2.74 mm thick with a 37.5 mm by 8.92 mm rectangular cross-section. The alumina/silica end pieces were nominally 2.79 mm thick with the same cross-section as the plates. The thickness of the glue layers was 0.036 mm. The mismatch between the shell and the stack was 0.254 mm. The shell's major and minor radii were 40.1 mm and 17.4 mm, respectively. The shell wall thickness was 4.6 mm, and a single shell's length was 48.3 mm. The transducer contained a pair of shell/stack assemblies. Since the two stacks were electrically staved together, we assumed that the effective length of the transducer was double the length of a single shell.

Table I lists the mechanical and thermal properties of the materials used in this transducer analysis. Table II lists the electromechanical parameters for composition 0.98(0.88PMN-0.12PT)-0.02BT as a function of tempera-

TABLE I. Mechanical and thermal properties for the transducer components.

Material	Modulus $1/s_{33}$ (GPa)	Density, ρ (kg/m^3)	Conductivity, λ W/(m K)	Specific heat, C_p [J/(kg K)]	Viscosity, μ (mPa s)
PMN-PT-BT	104.0	7900	3.59	350	-
Glue	1.0	1200	0.4	1700	-
Alumina/silica	64.1	2520	1.46	790	-
Titanium	115.0	4440	6.00	610	-
Neoprene		1.246	0.19	1900	-
SF_6 gas	-	6.140	0.014	666	0.015

TABLE II. Electromechanical properties for 0.98(0.88PMN-0.12PT)-0.02BT.

Temperature (°C)	Q_{33} (m^4/C^2)	P_s (C/m^2)	k (m/MV)	Tan δ
5	0.0152	0.264	2.860	0.132
10	0.0153	0.261	2.512	0.093
20	0.0151	0.251	1.951	0.038
30	0.0153	0.245	1.5967	0.013
40	0.0154	0.238	1.289	0.005
50	0.0156	0.229	1.078	0.003
60	0.0158	0.219	0.912	0.002
70	0.0162	0.207	0.785	0.002
80	0.0166	0.195	0.681	0.002

ture. We assumed in this model that the open-circuit compliance of PMN was independent of temperature.

Using the properties of SF_6 from Table I, we computed the heat transfer coefficient for the natural convection boundary condition. We determined that $Gr = 8.528 \times 10^6$, $Pr = 0.704$, $Nu = 31.88$, assuming a temperature difference of 45 °C between the heated ceramic and the gas. A heat transfer coefficient of $h = 12.26\ W/(m^2\,K)$ results from Eq. (9).

The transducer experiments in water were performed at 5 kHz. Figure 8 shows the predicted internal heat rate for the transducer unit with two shells and stacks. The applied ac voltage was 630 V and the dc bias voltage was 1900 V. The results also indicate a sharp drop in the heat rate between 20 and 40 °C. Figure 9 shows the source level drop from the 5 °C level predicted by the sonar model as a function of temperature. A sharp 12 dB drop exists over the 75 °C degree temperature span depicted in the figure. We also performed a simulation for the Navy PZT III with the same ac voltage using a linear piezoelectric constitutive law. That model predicted a source level 11 dB down from the PMN value at 5 °C, and Fig. 9 denotes that the result as a dashed line. The transducer's conductance and electrical power requirements also decreased with temperature.

Transient thermal computations were performed to predict the temperature rise for 7.5% and 20% duty cycles with a 630 V ac voltage and a 1900 V dc bias voltage. Figure 10 shows the predicted temperature as a function of time at the center of the ceramic stacks for the 7.5% duty cycle. The solid line represents the temperatures measured in the actual

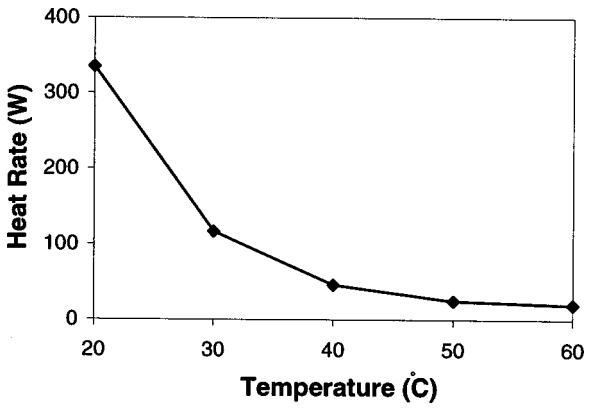


FIG. 8. Predicted heat rate as a function of temperature for a transducer unit operating at 5 kHz in water. The applied ac voltage was 630 V, and the dc bias voltage was 1900 V.

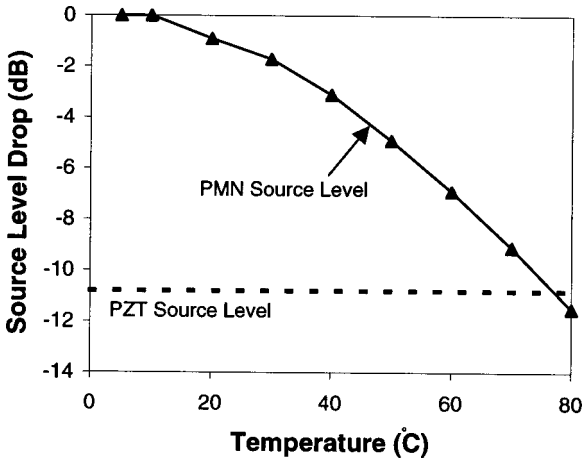


FIG. 9. The predicted source level for the flextensional transducer decreases dramatically with temperature. The applied ac voltage was 630 V, and the dc bias was 1900 V. The dashed line indicates the predicted source level for an identical flextensional transducer driven by Navy Type III PZT with the same voltage conditions.

experiment using thermocouples placed at the center. The initial temperature of the ceramic was 30 °C for both the model and the experiment. Figure 11 shows similar results for the 20% duty cycle. The starting temperature of the ceramic stack was 26 °C for both the model and the experiment.

As seen in Figs. 10 and 11, the thermal model predicts the initial ramp up and steady-state center temperature very well. The results show that the model underpredicted the steady-state temperature by about 3 °C for the 7.5% duty cycle, and predicted very accurately the steady-state temperature for the 20% duty cycle. For the lower duty cycle, the assumed convective heat transfer at the Neoprene-water interface was valid only for a short part of the duty cycle (when the ac voltage was on). When the ac voltage was off during the duty cycle, the correct mode of heat transfer was conduction through water and not convection. Therefore, the model underpredicted the peak temperature to a higher degree for lower duty cycle simulations.

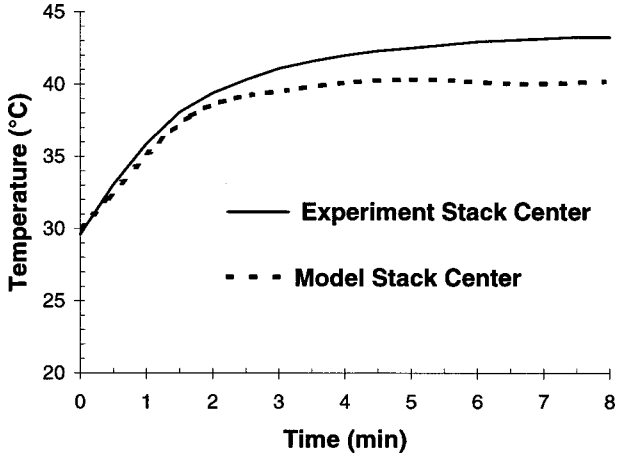


FIG. 10. Model prediction compared with experimentally measured temperature rise for the transducer in water with a 7.5% duty cycle, a 630 V ac voltage, and a 1900 V dc bias voltage. The initial temperature of the transducer was 30 °C.

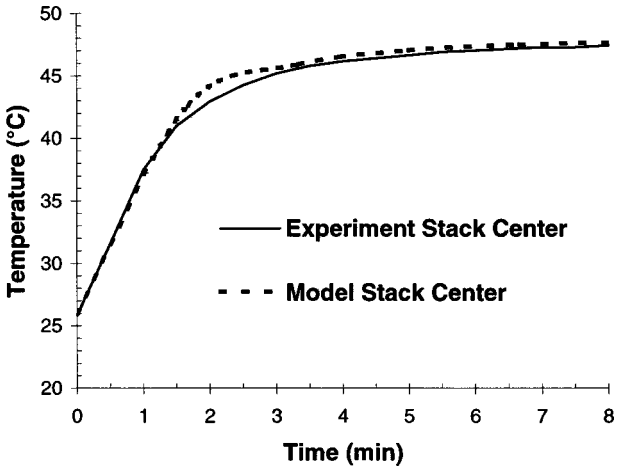


FIG. 11. Model prediction compared with experimentally measured temperature rise for the transducer in water with a 20% duty cycle, a 630 V ac voltage, and a 1900 V dc bias voltage. The initial temperature of the transducer was 26 °C.

Figure 12 shows the temperature distribution computed by the finite element analysis for the 20% duty cycle when thermal equilibrium was reached. The figure indicates that the generated heat primarily diffused by conduction from the stack's ends to the shell and neoprene. The PMN and the neoprene sleeve presented the main barrier to this conduction mechanism due to their low thermal conductivity. Consequently, a temperature gradient exists along the stack's length. The temperature at the edge of the ceramic at steady state was 38 °C compared to 47 °C at the center for the 20% duty cycle. A large temperature gradient also exists in the neoprene; the shell temperature for the 20% duty cycle was 25 °C above the ambient water temperature.

After the thermal model reached equilibrium, we predicted the relative drop in source level using the average temperature of the PMN stacks and the acoustic results displayed in Fig. 9. For the 7.5% duty cycle, the model predicted an average ceramic temperature of 34 °C and a 2.3 dB drop in source level relative to the value at 5 °C. For the 20% duty cycle, we predicted an average ceramic temperature of 41 °C and a 3.2 dB drop in source level relative to the 5 °C value. The steady state source level measured for the 20% duty cycle was 2 dB higher than the predicted value. This result indicates that the acoustic model tends to underpredict the source level. Hom and Shankar[1] observed the same underprediction with their isothermal model of an electrostrictive flextensional transducer.

IV. CONCLUSIONS

The numerical and experimental results presented in this paper show that a significant temperature rise occurred in the PMN ceramic due to hysteretic heating. The temperature increase was over 40 °C above the ambient temperature of the surrounding water. The PMN composition used had a relatively high Curie transition temperature of 45 °C (as measured with a 1 kHz weak field voltage source). This high transition material was selected to reduce the source level drop at the equilibrium temperature. While our predictions show that a 3 dB drop in source level occurred, the predicted steady-state value was still nearly 8 dB higher than that predicted for an identical PZT transducer with the same voltage conditions. We also obtained good agreement between the model predictions and experiments. The model matched the steady-state temperature: underestimating for one set of conditions by only 3 °C and estimating to within 1 °C for the other. The model accurately predicted the transient temperature rise rate in both cases as demonstrated by Figs. 10 and 11. The final source level predicted by the model was also within 2 dB of the measured source level for both cases.

V. DISCUSSION

As noted in Sec. I, we made an approximation in the coupling of our acoustic model with our thermal model; we performed the sonar computations using an average ceramic temperature. This assumes that the mean temperature, not a temperature distribution, characterized both the acoustic source level and heat generation rate. We subsequently used the heat generation rates in the thermal analysis of the device's temperature distribution and history. This assumption loosely couples the two models and makes a solution tractable. We validated this approximation by comparing our results with experimental measurements on the exact system modeled. We assert that the very good model/experiment match for two different sets of starting conditions and duty cycles demonstrates that our initial approximation was adequate.

An alternative validation approach would match our predicted results with those of a more sophisticated model, which does not make our approximation. This model would fully couple the acoustic, structural, electrical, and thermal degrees of freedom. We believe that this type model is not feasible for the specific problem studied in this paper, since the acoustic and thermal responses occur over dissimilar time scales. The acoustics behavior occurs at 1/5000 of a

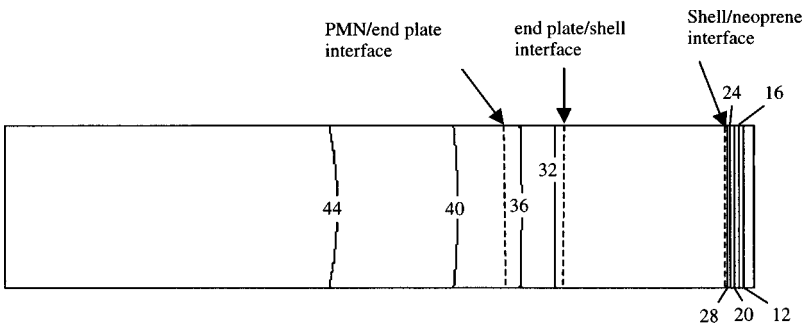


FIG. 12. Temperature distribution in °C at equilibrium as predicted by finite element analysis for a 20% duty cycle starting at 26 °C.

second (5 kHz), while the thermal response requires over 8 min to reach full steady state. Assuming 15 steps per cycle for explicit time integration, this means 36 million time steps to model the full transient response, which exceeds current computational technology. At best, this method could only improve the predicted steady-state temperature by 3 °C over our model.

The acoustic/thermal model also ignored the temperature dependence of the glue layers that bonded the PMN plates. For some types of PZT transducers experiencing thermal runaway, melting glue layers could significantly reduce the ceramic stack stiffness. The epoxy resin system used for the transducer presented in this paper had an 80 °C cure temperature, which the transducer never exceeded during operation. While this transducer experienced a large temperature rise, it was not hot enough to melt the bond joints.

We can make improvements in our model to increase the accuracy of the acoustic and thermal predictions. First, the PMN material properties used for our analysis were based on quasi-static measurements at 1 Hz with a high electric field. Because it's a relaxor ferroelectric, PMN's dielectric permittivity and hysteresis depend strongly on frequency.[3] Both of these properties affect the heating rate in the ceramic. More nearly accurate simulations using our model will require material parameters based on measurements made with both high electric fields and frequencies. While these conditions would accurately simulate the material state in an operational transducer, the experimental setup would require high-powered amplifiers and proper voltage protection.

As a design tool, our acoustic/thermal model could also investigate different compositions of PMN to predict their corresponding steady-state acoustic output before building a transducer. Characterizing the electromechanical behavior of a material only requires a small number of inexpensive experiments, while characterizing a transducer requires expensive construction and acoustic testing. Our model would allow screening of new PMN variants early in the materials development process to determine the optimum composition and Curie temperature. This design approach could accelerate insertion of new PMN materials into sonar systems. The model could also investigate improved PMN transducer designs incorporating active or passive cooling techniques. An inert liquid could fill the shell cavity to create convection rolls, or heat exchangers could remove heat from the ceramic.

ACKNOWLEDGMENTS

The authors gratefully acknowledge helpful discussions with Dr. Roger Richards of the Naval Undersea Warfare Center; Dr. Keith Bridger of Active Signal Technologies; Dr. Steve Winzer and Ira Chapman of Lockheed Martin Missiles & Space; and Jerome DeJaco and David Gillette of the Space & Naval Warfare Systems Command. The Naval Undersea Warfare Center and Active Signal Technologies measured the transducer heating data. Pamela Sadler-Hom and Edger Seydel of Lockheed Martin Missiles & Space made the electromechanical measurements on the PMN material. The authors would like to thank David Gillette for providing the CHIEF code. The Office of Naval Research (ONR) funded this effort. Dr. Ken Dial was the ONR Program Officer, and Greta Conde was the PEO (USW) Program Manager.

[1] C. L. Hom and N. Shankar, ''Modeling nonlinearity in electrostrictive sonar transducers,'' J. Acoust. Soc. Am. **104**, 1903–1913 (1998).
[2] S. M. Pilgrim, M. Massuda, J. D. Prodey, and A. P. Ritter, ''Electrostrictive sonar drivers for flextensional transducers,'' in *Transducers for Sonics and Ultrasonics*, edited by M. McCollum, B. F. Hamonic, and O. B. Wilson (Technomic, Lancaster, PA, 1993).
[3] S. M. Pilgrim, M. Massuda, J. D. Prodey, and A. P. Ritter, ''Electromechanical properties of some $Pb(Mg_{1/3}Nb_{2/3})O_3$–$PbTiO_3$–(Ba, Sr)TiO_3 ceramics: Part I,'' J. Am. Ceram. Soc. **75**, 1964–1969 (1992).
[4] C. L. Hom and N. Shankar, ''A fully coupled constitutive model for electrostrictive ceramic materials,'' J. Intell. Mater. Syst. Struct. **5**, 795–801 (1994).
[5] H. A. Schenck, ''Improved integral formulation for acoustic radiation problems,'' J. Acoust. Soc. Am. **44**, 41–58 (1967).
[6] G. W. Benthien, D. Barach, and D. Gillette, CHIEF User Manual, Report No. TD970 (NOSC, San Diego, CA, 1988).
[7] ABAQUS, Version 5.8 (Hibbitt, Karlsson and Sorensen, Inc., Pawtucket, RI, 1998).
[8] J. L. Butler, T. J. Peirce, and J. Lindberg, ''A desktop computer program for a flextensional transducer,'' Proceed. Instit. Acoust. **9**, 31–41 (1987).
[9] G. E. Martin, ''On the theory of segmented electromechanical systems,'' J. Acoust. Soc. Am. **36**, 1366–1370 (1964).

Novel graph model and analysis method for piezoelectric thickness-drive transducers

Wen-Jeng Hsueh[a)]
Department of Naval Architecture and Ocean Engineering, National Taiwan University, Taiwan, Republic of China

(Received 15 November 1999; revised 18 April 2000; accepted 26 July 2000)

This study develops a new graph model and analysis scheme for representing and analyzing a piezoelectric thickness-drive transducer. The proposed TWSF graph model, expressed via electrical and acoustical variables instead of electromechanical analogy, simplifies understanding of the physical mechanism of the transducer. The primary advantage of the graph model analogy is that performance characteristics, such as the electrical and mechanical impedance, and the force and velocity responses of the transducer system can be calculated directly using the topology scheme based on the graph model. The dynamic equations do not need to be formulated according to the analogy model for design and analysis. The proposed method also does not require computation schemes to solve the equations for obtaining the characteristics and the responses. Compared to the classical method, the present graph and analysis scheme offers a more efficient technology to obtain precise results with less analysis. Moreover, the derived results are expressed as both analytical and closed-form formulas, making them convenient to apply in analysis and design. © *2000 Acoustical Society of America.* [S0001-4966(00)00511-7]

PACS numbers: 43.38.Fx, 43.20.Wd [SLE]

I. INTRODUCTION

Piezoelectric thickness-drive transducers have been widely applied in various fields, such as underwater communication and detection, nondestructive evaluation, and medical diagnosis. The dynamic behavior of the piezoelectric transducer includes the electrical and acoustical mechanisms, and their interaction is quite complex. Understanding these mechanisms and simplifying the analysis are the two major objectives of modeling the transducers. Mason[1] depicted the piezoelectric thickness-drive transducer as an electrical analogy model. Later, Redwood,[2] and Krimholtz *et al.*[3] proposed a few variations of this model. Since a negative capacitance and two trigonometric elements are used, these nonstandard circuit elements limit the application of standard analysis tools for the network in the analysis. Schwarz,[4] Morris and Hutchens,[5] and Leach, Jr.[6] presented the use of the numerical scheme or computer package for the computation. Meanwhile, Busch-Vishniac and Paynter[7] discussed other disadvantages of the electrical analogy. Thus Hayward and Jackson[8] proposed block diagram models to express the relationship of the force transmission and the electrical load. Moon and Busch-Vishniac[9] used a bond graph model to describe the power flow transmission. Although such models can provide good physical realizability, their use to reduce analysis work is limited.

Generally, engineers apply these classical models to formulate the dynamic equations of the combined transducer systems, including the front and back loads. Since the formulated equations are coupled, it is difficult to obtain the analytical solution of the system characteristics and dynamic responses, such as the input electric impedance, force, and velocity. Hence, various numerical schemes, such as the matrix methods, are applied to compute the results by Allik and Huges,[10] Hayward and Jackson,[8] Moon and Busch-Vishniac,[9] and Kraszewski and Sigelmann.[11]

This study develops a TWSF (two-way state-flow) graph model and analysis method for the piezoelectric thickness-drive transducer to cut down analysis. The author[12,13] has developed various TWSF graph models to investigate vibration and wave propagation in the lumped and continuous systems. The advantage of using the graph model for the transducer analysis is that the system characteristics and responses can be directly calculated from the graph model via the topology scheme. Thus formulating the dynamic equations according the analogy model is unnecessary. Meanwhile, using computation schemes to solve the equations for obtaining the characteristics and responses is also not required. Moreover, the derived results, including the characteristics and responses, are expressed as analytical and closed forms. These expressions provide further insight into how the individual variables influence the desired results. This additional advantage is vital to good synthesis and design.

The rest of the paper is organized as follows. Section II develops TWSF graph model of the transducer. Meanwhile, Sec. III derives analytical system characteristics and responses, and Sec. IV compares the presented method with the classical method. Finally, Sec. V investigates examples of a PZT 5A piezoelectric transducer.

II. DEVELOPMENT OF THE TWSF GRAPH MODEL

For a piezoelectric transducer, if the magnetic and pyroelectric effects are neglected, the piezoelectric equations for the stress T_i and electric field E_m are given as[1,14] (e.g., Mason, 1964; Wilson, 1988):

a)Electronic mail: hsuehwj@ccms.ntu.edu.tw

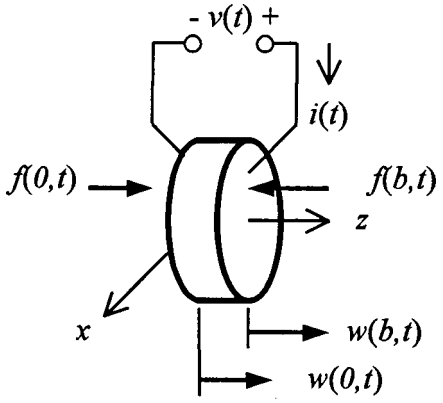


FIG. 1. Schematic diagram of the piezoelectric thickness-drive transducer.

$$T_i=\sum_{j=1}^{6} c_{ij}^{D}S_j-\sum_{k=1}^{3} h_{ik}D_k, \tag{1}$$

$$E_m=-\sum_{j=1}^{6} h_{jm}S_j+\sum_{k=1}^{3} \beta_{mk}^{S}D_k, \tag{2}$$

where c_{ij}^{D} is the elastic stiffness, h_{ik} is the piezoelectric constant, β_{mk}^{S} is the dielectric impermeability constant, S_i is the strain component, and D_i is the electric displacement. For a piezoelectric transducer with vibrations in the thickness directions as shown in Fig. 1, the strains in the transverse directions are assumed to be neglected. The only nonzero strain in S_3. Since the thickness direction is the poled direction, the only electric displacement component is D_3. Thus the piezoelectric equations can be simplified as

$$T_3=c_{33}^{D}S_3-h_{33}D_3, \tag{3}$$

$$E_3=-h_{33}S_3+\beta_{33}^{S}D_3. \tag{4}$$

According to the properties of an elastic and continuous material, the differential equation of the piezoelectric transducer in the thickness direction can be expressed as

$$c_{33}^{D}\frac{\partial^2 w(z,t)}{\partial z^2}-\rho\frac{\partial^2 w(z,t)}{\partial t^2}=0, \tag{5}$$

where ρ is the mass density of the transducer. When a sinusoidal excitation is applied, the general solution of the displacement is given as

$$w(z,t)=\left(\frac{U_1}{j\omega}\cos kz+\frac{U_2-U_1\cos kb}{j\omega\sin kb}\sin kz\right)e^{j\omega t}, \tag{6}$$

where k is the wave number in the transducer, defined by $k=\omega/c_t^{D}$, c_t^{D} is the phase speed of wave propagation $c_t^{D}=(c_{33}^{D}/\rho)^{1/2}$, and U_1 and U_2 are the phasor velocities of the surface at $z=0$ and $z=b$, respectively. The reflection force $f(z,t)$ acting on the surface plane at $z=0$ and $z=b$ should be also fitted to the stress of the transducer on both boundaries. Substituting Eq. (6) into the strain term S_3 of Eq. (3) gives

$$F_2=\frac{1}{\cos kb}(F_1-A_0h_{33}D_0)+\frac{A_0c_{33}^{D}k\tan kb}{j\omega}U_2+A_0h_{33}D_0, \tag{7}$$

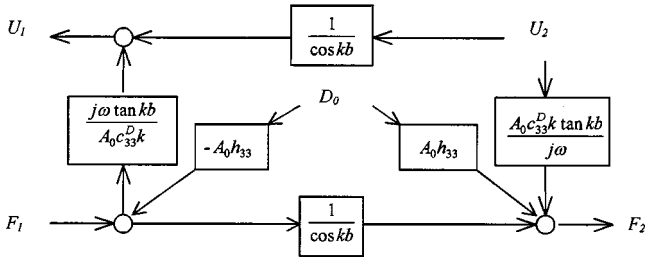


FIG. 2. Graph of Eqs. (7) and (8).

$$U_1=\frac{1}{\cos kb}U_2+\frac{j\omega\tan kb}{A_0c_{33}^{D}k}(F_1-A_0h_{33}D_0), \tag{8}$$

where A_0 is the surface area of the vibrator. F_1 and F_2 are the phasor forces on the surface of the transducer at $z=0$ and $z=b$, respectively, and D_0 is the phasor electric displacement in the z direction. Equations (7) and (8) can be expressed graphically, as shown in Fig. 2. The circle described in the graph is the node which represents the sum of states coming into it. Each state outward from the node is the same as the state of the node.

Let the voltage difference applied on both surfaces of the transducer be $v(t)$, and the relationship between the applied voltage and the electric field is given by

$$v(t)=\int_0^b E_3\,dz. \tag{9}$$

By substituting Eq. (4) into Eq. (9), the electric displacement can be expressed by the velocities of both surfaces of the transducer, based on Eqs. (3) and (6), as

$$D_0=\frac{1}{\beta_{33}^{S}b}V+\frac{h_{33}}{j\omega\beta_{33}^{S}b}(U_2-U_1), \tag{10}$$

where V is the phasor voltage.

From the definition, the induced current $i(t)$ is equal to the area times derivative of the electric displacement in the z direction with respect to time. Thus the relation between the induced phasor current I and the phasor electric displacement is given as

$$I=j\omega A_0D_0. \tag{11}$$

Equations (10) and (11) are described as Fig. 3.

Figures 2 and 3 can be combined as a graph model for the transducer as shown in the middle of Fig. 4. In the graph model, the electrical current I, the force F_2 and velocity U_1 are all functions of the electrical voltage V, the force F_1 and velocity U_2. From the graph model, we can observe the conversion of the electrical and the mechanical states of the transducer and the loads.

In the graph model, the forces at the left and right terminals are all in the lower horizontal path and in the same direction from left to right. On the other hand, both the velocities are arranged in the upper horizontal path and in the same direction from right to left. This special arrangement of the graph representation can provide convenience in connecting the graph model of the loads, backing, transformer and the transducer for analysis.

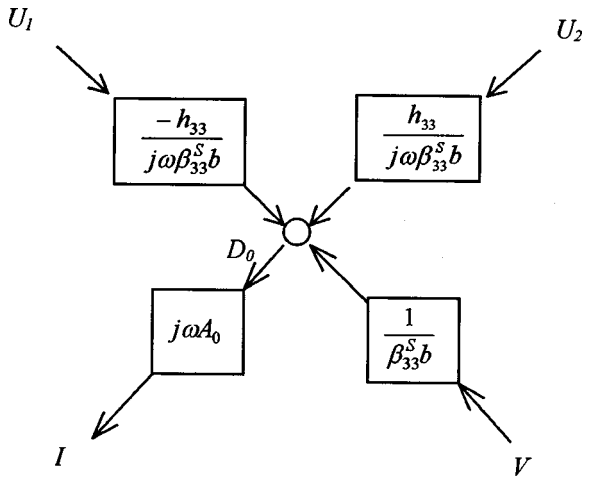


FIG. 3. Graph of Eqs. (10) and (11).

III. PERFORMANCE ANALYSIS

A. Electrical impedance and admittance

In practical applications, the front and back faces of the transducer are usually connected to external loads. If the front and back loads, with the mechanical admittance Y_2 and impedance Z_1, respectively, are considered in coupling the transducer, both impedance and admittance seen from the transducer can be cascaded to the terminal of the graph model of the transducer to form the closed-form model as shown in Fig. 4.

Based on the graph model developed, the transfer characteristics and impedance can be calculated using the topology formula (Mason, 1956)[15] described as

$$H=\sum_i P_i B_i/D, \tag{12}$$

where H is the phasor of the response, P_i is the path gain of the ith forward path, D is the determinant of the graph, and B_i is the cofactor of the ith forward path determinant of the graph with the loops touching the ith forward path removed.

In Fig. 4, we see that there are eight loops in the graph model. The flow path of each loop is listed in Table I. For example, the path of loop 1 is beginning from node a through the gain $-Z_1$, the node b, and return to a. The loop gain is the multiplication of each gain which the path of the loop passes. The loop gain L_i for each loop i is given as

TABLE I. State-flow path and the nodes it passes in each closed loop.

Loop number	Loop gain	The nodes passed by the path of the loop
1	L_1	$a\rightarrow(-Z_1)\rightarrow b\rightarrow a$
2	L_2	$a\rightarrow(-Z_1)\rightarrow b\rightarrow c\rightarrow(Y_2)\rightarrow d\rightarrow a$
3	L_3	$a\rightarrow e\rightarrow b\rightarrow a$
4	L_4	$e\rightarrow b\rightarrow c\rightarrow(Y_2)\rightarrow d\rightarrow e$
5	L_5	$e\rightarrow b\rightarrow c\rightarrow(Y_2)\rightarrow d\rightarrow a\rightarrow e$
6	L_6	$e\rightarrow c\rightarrow(Y_2)\rightarrow d\rightarrow e$
7	L_7	$e\rightarrow c\rightarrow(Y_2)\rightarrow d\rightarrow a\rightarrow e$
8	L_8	$c\rightarrow(Y_2)\rightarrow d\rightarrow c$

$$L_1=-\frac{j\omega\tan kb}{A_0 c_{33}^D k}Z_1, \tag{13}$$

$$L_2=-\frac{Z_1 Y_2}{\cos^2 kb}, \tag{14}$$

$$L_3=\frac{h_{33}^2\tan kb}{\beta_{33}^S b c_{33}^D k}, \tag{15}$$

$$L_4=\frac{-A_0 h_{33}^2}{j\omega\beta_{33}^S b\cos kb}Y_2, \tag{16}$$

$$L_5=\frac{A_0 h_{33}^2}{j\omega\beta_{33}^S b\cos^2 kb}Y_2, \tag{17}$$

$$L_6=\frac{A_0 h_{33}^2}{j\omega\beta_{33}^S b}Y_2, \tag{18}$$

$$L_7=L_4, \tag{19}$$

$$L_8=\frac{A_0 c_{33}^D k\tan kb}{j\omega}Y_2. \tag{20}$$

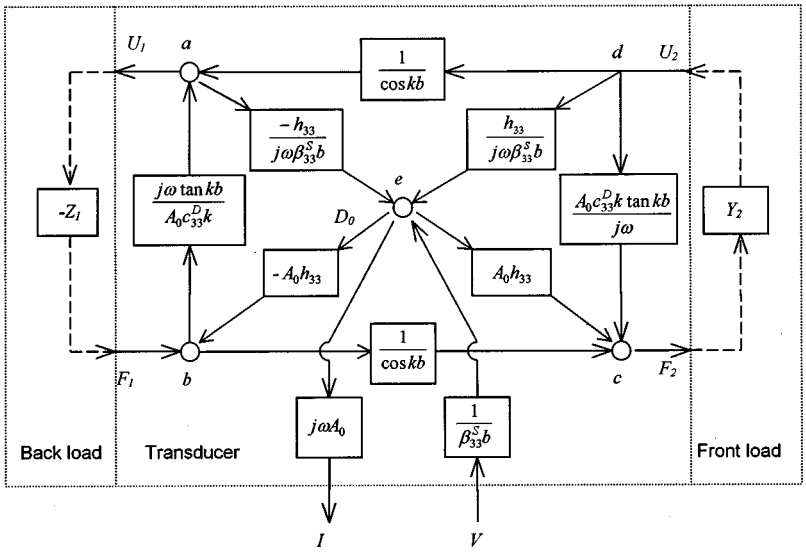


FIG. 4. TWSF graph model of the thickness-drive transducer connecting loads.

TABLE II. State-flow path and the nodes it passes in each forward path.

Forward path	The nodes passed by the forward path
From V to I	Path 1: $V\rightarrow e\rightarrow I$
From V to F_2	Path 1: $V\rightarrow e\rightarrow b\rightarrow c\rightarrow F_2$ Path 2: $V\rightarrow e\rightarrow c\rightarrow F_2$
From V to F_1	Path 1: $V\rightarrow e\rightarrow b\rightarrow a\rightarrow(-Z_1)\rightarrow F_1$ Path 2: $V\rightarrow e\rightarrow c\rightarrow(Y_2)\rightarrow d\rightarrow a\rightarrow(-Z_1)\rightarrow F_1$ Path 3: $V\rightarrow e\rightarrow b\rightarrow c\rightarrow(Y_2)\rightarrow d\rightarrow a\rightarrow(-Z_1)\rightarrow F_1$

Of these eight loops, there are only three pairs of two nontouching loops, L_1 and L_6, L_1 and L_8, and L_3 and L_8. According to the determinant defined previously, the determinant of the graph model can be expressed by

$$D=1-\sum_{i=1}^{8} L_i+L_1L_6+L_1L_8+L_3L_8. \tag{21}$$

Substituting Eqs. (13)–(20) into Eq. (21), the determinant leads to

$$D=\frac{j\omega\tan kb}{A_0c_{33}^Dk}Z_1-\frac{c_{33}^Dk\tan kb}{j\omega}\left(1-\frac{\alpha\tan(kb/2)}{kb/2}\right)Y_2+\left(1-\frac{\alpha\tan kb}{kb}\right)(1+Z_1Y_2), \tag{22}$$

where α is a nondimensional constant defined by $h_{33}^2/(\beta_{33}^Sc_{33}^D)$.

If the electrical admittance of the transducer is to be calculated, according to the graph model, there is only one forward path from V to I with the path gain of $j\omega A_0/(\beta_{33}^sb)$. The nodes passed by the forward path are illustrated in Table II. As seen in Fig. 4, only loops 1, 2, and 8 do not touch the forward path. From the formula Eq. (12), the electrical admittance can be obtained as

$$Y_p=j\omega A_0(1-(L_1+L_2+L_8)+L_1L_8)/(\beta_{33}^sbD). \tag{23}$$

Since the admittance is the inverse of the impedance, it is easy to acquire the electrical impedance given as

$$Z_p=\frac{\beta_{33}^sbc_{33}^Dk D}{j\omega A_0c_{33}^Dk(1-Y_2Z_1)+\tan(kb)(\omega^2Z_1-(A_0c_{33}^Dk)^2Y_2)}. \tag{24}$$

B. Force output response

For calculating the response of the force output from the front surface of transducer F_2 with respect to the electric voltage input V, two forward paths are there to follow the direction of the state flow in the graph model from V to F_2 displayed in Table II. One of the forward paths, with path gain $-A_0h_{33}/(\beta_{33}^sb\cos kb)$, touches all the loops while the other with path gain $A_0h_{33}/(\beta_{33}^sb)$, touches all the loops except loop 1. Thus the response can be calculated as

$$\frac{F_2}{V}=\frac{A_0h_{33}}{\beta_{33}^SbD}\left(1-L_1-\frac{1}{\cos kb}\right). \tag{25}$$

Substituting Eq. (13) into Eq. (25) gives

$$F_2=\frac{h_{33}\tan(kb)(j\omega Z_1-A_0c_{33}^Dk\tan(kb/2))V}{\beta_{33}^Sc_{33}^DkbD}. \tag{26}$$

To calculate the response of the force output from the back surface, there are three forward paths from V to F_1, illustrated in Table II. The first path, with path gain $j\omega h_{33}Z_1\tan kb/(\beta_{33}^sc_{33}^Dkb)$, does not pass through load 2, but the others, with path gains $-A_0h_{33}Y_2Z_1/(\beta_{33}^sb\cos kb)$ and $A_0h_{33}Y_2Z_1/(\beta_{33}^sb\cos^2 kb)$, do pass through the load. The first forward path touches all the paths except for loop 8. However, the other two forward paths touch all the loops of the graph model. Thus the response can be expressed as

$$\frac{F_1}{V}=\frac{j\omega h_{33}Z_1\tan kb}{\beta_{33}^Sc_{33}^DkbD}(1-L_8)+\frac{A_0h_{33}Y_2Z_1}{\beta_{33}^sbD\cos kb}\times\left(-1+\frac{1}{\cos kb}\right). \tag{27}$$

Substituting Eq. (20) into Eq. (27) gives

$$F_1=\frac{h_{33}\tan(kb)(j\omega-c_{33}^Dk A_0Y_2\tan(kb/2))Z_1V}{\beta_{33}^Sc_{33}^DkbD}. \tag{28}$$

C. Front and back velocity

Since the impedance and admittance of the load are given as ratios of force and velocity at the surfaces of the transducer, the relationships between the front and back velocity and the force are

$$U_1=\frac{-F_1}{Z_1}, \tag{29}$$

$$U_2=F_2Y_2. \tag{30}$$

According to Eqs. (29) and (30), the front and back velocity can be calculated if forces F_1 and F_2 have been obtained by Eqs. (26) and (28).

D. Mechanical loads included

In many applications, a backing is connected to the back face of the transducer, and a transformer (matching layers) is inserted between the front surface of the transducer and the front loads. The TWSF graph models of the backing and the transformer can be obtained by just neglecting the electric effect from the graph model of the piezoelectric transducer as Fig. 4 depicts. The graph models for both of the transformer and the front load, and both the backing and the back load can be combined to equivalent impedance and admittance, respectively, using the model reduction scheme.[13]

E. Electrical source included

When the piezoelectric transducer connected to the terminating electrical circuit is considered, the circuit can be expressed as the Thevenin equivalent circuit, as shown in Fig. 5(a). According to the circuit theory, the relationship between the equivalent voltage source and the voltage of the input terminal of the transducer are expressed as

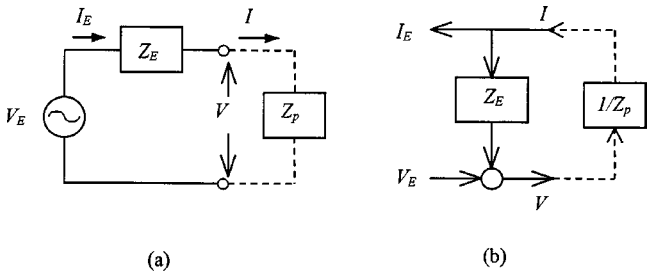


FIG. 5. Terminating electric circuit, (a) Thevenin equivalent circuit, (b) TWSF graph model.

$$V = V_E - I_E Z_E, \tag{31}$$

$$I = I_E, \tag{32}$$

where V_E and Z_E are the Thevenin equivalent voltage and equivalent impedance of the terminating electrical circuit, respectively. Equations (31) and (32) can be expressed as a graph model, as shown in Fig. 5(b). From Figs. 4 and 5(b) we see that the graph model of the terminating electrical circuit just fits that of the piezoelectric transducer. Thus the terminating electrical circuit can be easily included in the model for analysis.

IV. COMPARISON WITH CLASSICAL MODELS

To illustrate the performance difference between using the TWSF graph model and classical models for the design and analysis of the piezoelectric transducers, this study compares the design procedures using these models. The classical equivalent circuit model[1,15] can be constructed as Fig. 6 indicates. Meanwhile, the electrical components and parameters of the equivalent circuit are expressed as

$$Z_{E1} = jA_0 \rho c_t^D \tan\frac{kb}{2}, \tag{33}$$

$$Z_{E2} = \frac{-jA_0 \rho c_t^D}{\sin kb}, \tag{34}$$

$$C_0 = \frac{A_0}{\beta_{33}^S b}, \tag{35}$$

$$N = C_0 h_{33}. \tag{36}$$

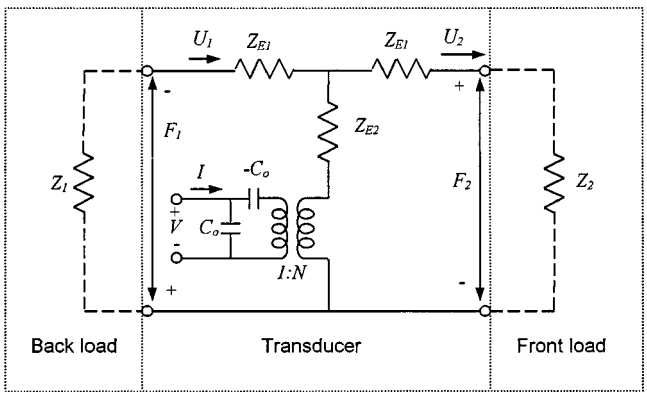


FIG. 6. Equivalent circuit model of the thickness-drive transducer.

In the design and analysis of a transducer system, the equivalent circuit model is given first. Generally, the dynamic equations of the transducer can be formulated using the circuit network theory, and are usually expressed in matrix form.

$$\begin{Bmatrix} F_1 \\ F_2 \\ V \end{Bmatrix} = \begin{bmatrix} -(Z_{E1+Z_{E2}}) & Z_{E2} & \dfrac{-N}{j\omega C_0} \\ Z_{E2} & -(Z_{E1}+Z_{E2}) & \dfrac{N}{j\omega C_0} \\ \dfrac{N}{j\omega C_0} & \dfrac{-N}{j\omega C_0} & \dfrac{1}{j\omega C_0} \end{bmatrix} \begin{Bmatrix} U_1 \\ U_2 \\ I \end{Bmatrix}. \tag{37}$$

When the front and back loads connected to the transducer are considered, both loads can be modeled as the equivalent impedances Z_2 and Z_1 connected to the terminals of the equivalent circuit model, respectively. Thus the relationship between U_1 and F_1 depends on Z_1, and that between U_2 and F_2 depends on Z_2. Both relations can be used to formulate two additional dynamic equations. In this way the system performance and the force and velocity responses can be simulated numerically.[8,11]

When the backing and transformer are considered, their equivalent circuit models are generally expressed as a T network,[1,14,16] which resembles the circuit model of the transducer excluding the transformer. Since the terminals of the equivalent circuit models of the backing and transformer coincide with those of the transducer, those circuit models can be directly connected in series. Meanwhile, additional dynamic equations can be formulated by the circuit network theory. Thus the matrix method[8,11] and computer software[6] are normally applied to the numerical simulation of the system performance, force, and velocity.

The typical design and analysis procedures for a piezoelectric transducer system are based on the classical models, including the equivalent circuit and bond graph models, and the TWSF graph method is summarized as below.

A. Design procedures based on the equivalent circuit model

(1) Sketching the equivalent circuit model of the transducer system, including the piezoelectric transducer, the transformers, and the front and back loads.
(2) Formulating the dynamic equations based on the circuit model using circuit theory (KVL or KCL).
(3) Computing numerical simulation of the force and velocity response as well as system characteristics via solving the formulated dynamic equations.

B. Design procedures based on the bond graph model

(1) Sketching the bond graph model of the transducer system, including the piezoelectric transducer, transformers, and front and back loads.
(2) Formulating the dynamic equations based on the analogy model by bond graph notation.

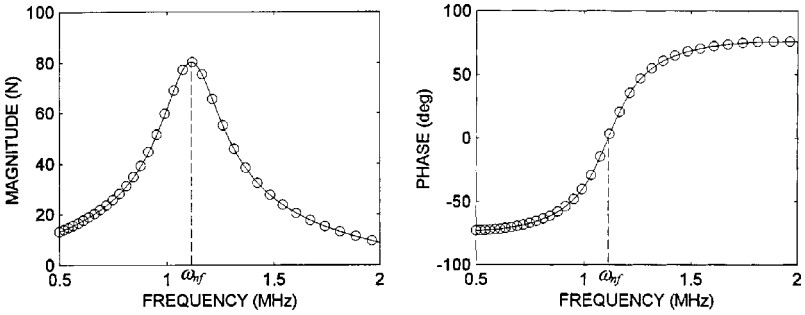


FIG. 7. Force output response of the piezoelectric transducer for input voltage of 100 V. Solid line: the graph method; ○: the equivalent circuit method.

(3) Computing numerical simulation of the force and velocity response as well as system characteristics via solving the formulated dynamic equations.

C. Design procedures based on the TWSF graph model

(1) Sketching the TWSF graph model of the transducer system, including the piezoelectric transducer, transformers, and front and back loads.
(2) Calculating the analytical results of the force and velocity response and the system characteristics via graph theory directly from the graph model.

The above design procedures reveal that the third step of the design procedures based on the equivalent circuit and bond graph models is the most difficult and time consuming. Thus related investigations[4–6] presented various advanced numerical schemes and software to simplify the computation. However, the design procedures based on the TWSF model avoid the tedious third step. This indicates that the new graph model can allow more efficient analysis and design of the transducer system than classical models.

V. EXAMPLES

To illustrate the implementation of the proposed model, a PZT 5A thickness-drive transducer with area $A_0=3.14\times10^{-4}\,\mathrm{m}^2$ and thickness $b=1.74\,\mathrm{mm}$ is investigated by the present and classical methods. The load on the front and back of the transducer is assumed to be water with mechanical impedance $Z_2=A_0\times(1.5\times10^6\,\mathrm{Pa\,s/m})$ and epoxy resin with mechanical impedance $Z_1=A_0\times(9.1\times10^6\,\mathrm{Pa\,s/m})$, respectively. Force output response from the front surface with respect to electric voltage input 100 V and the electrical impedance of the piezoelectric transducer are calculated using the present method and classical method, in which Eqs. (37) and (24), obtained by the TWSF graph model, are used for the analysis in the present method. In the classical method, the dynamic equation (26) formulated the equivalent circuit model, and the numerical scheme by matrix method is applied to solve the responses.

The force output from the front surface with respect to electric voltage input 100 V is first studied. Figure 7(a) displays that the frequency response near the first resonance mode calculated by both methods is closely matched. The resonance of the force output is at the first resonance frequency of 1.113 MHz, denoted as ω_{nf}, with an amplitude of 80.3 N. The phase of the force response calculated by both methods are identical illustrated in Fig. 7(b).

In the second case, the electrical impedance of the transducer is considered. Figure 8 shows that the antiresonance frequency of the electrical impedance is 1.299 MHz, denoted as ω_{ni}, at which the magnitude is 141.9 N/V. From the figure, we see that the resonance frequency and the magnitude calculated by both the present method and the equivalent circuit methods are the same. According to Figs. 7 and 8, we also see that the resonance frequencies of the electrical impedance and the force output are not identical. The difference between the two resonance frequencies is 0.064 MHz (5.7%).

According to the previous studies, we see that the force response and electrical impedance calculated by the present method are almost equal to those by the classical method over a wide frequency range.

VI. CONCLUSIONS

This investigation has created a TWSF graph model of a piezoelectric thickness-drive transducer. Based on the proposed model, performance characteristics, including electrical impedance and the force and velocity output responses, are derived via the topology scheme.

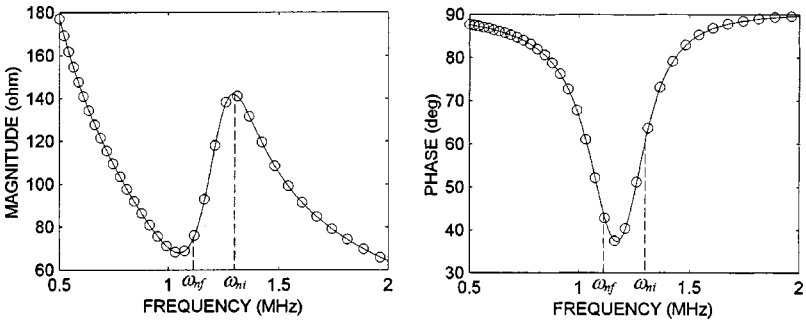


FIG. 8. Electrical impedance of the piezoelectric transducer. Solid line: the graph method; ○: the equivalent circuit method.

Compared to classical methods and models for the analysis of transducer systems, the graph model developed herein has the following advantages:

(1) The system characteristics and dynamic responses of the transducer system can be calculated directly from the graph model. Formulating the dynamic equations and solving the results by numerical technologies is unnecessary.
(2) The results derived from the presented method are expressed in analytical and closed forms, improving intuitive understanding of how each variable influences the results.
(3) Drawing the TWSF graph model according to the basic governing equations of the transducers is convenient. This representation offers better qualitative and quantitative physical insights. Moreover, understanding the bond graph and equivalent circuit analogy is not required.

Finally, numerical examples of transducer analysis using the proposed graph model and analysis scheme and the classical method based on the equivalent circuit model have confirmed that the present model and analysis method involve less analysis than other methods.

Although this study has only examined the thickness-drive transducer, the method could potentially be extended to other types of piezoelectric transducers.

ACKNOWLEDGMENT

This research was supported in part by the National Science Council of the Republic of China under Grant No. NSC 88-2611-E-002-009.

[1] W. P. Mason, *Physical Acoustics*, Vol. I, Part A (Academic, New York, 1964).
[2] M. Redwood, ''Transient performance of a piezoelectric transducer,'' J. Acoust. Soc. Am. **33**, 527–536 (1961).
[3] R. Krimholtz, D. A. Leedom, and G. L. Matthaei, ''New equivalent circuit for elementary piezoelectric transducers,'' Electron. Lett. **6**, 338–339 (1970).
[4] R. Schwarz, ''Digital computer simulation of a piezoelectric thickness vibrator,'' J. Acoust. Soc. Am. **62**, 463–367 (1977).
[5] S. A. Morris and C. G. Hutchens, ''Implementation of Mason's model on circuit analysis programs,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **33**, 295–298 (1986).
[6] W. M. Leach, Jr., ''Controlled-source analogous circuits and SPICE models for piezoelectric transducers,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **41**, 60–66 (1994).
[7] I. J. Busch-Vishniac and H. M. Paynter, ''Bond graph models of sound and vibration systems,'' J. Acoust. Soc. Am. **85**, 1750–1758 (1989).
[8] G. Hayward and M. N. Jackson, ''A lattice model of the thickness-mode piezoelectric transducer,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **33**, 41–50 (1986).
[9] W. K. Moon and I. J. Busch-Vishniac, ''A finite-element equivalent bond graph modeling approach with application to the piezoelectric thickness-mode transducer,'' J. Acoust. Soc. Am. **93**, 3496–3506 (1993).
[10] H. Allik and T. Huges, ''Finite element methods for piezoelectric vibration,'' Int. J. Num. Methods Eng. **2**, 151–157 (1970).
[11] R. J. Kraszewski and R. A. Sigelmann, ''A chain matrix model for an acoustically coupled, lead zirconate titanate core, electrical transformer,'' J. Acoust. Soc. Am. **90**, 2763–2768 (1991).
[12] W. J. Hsueh, ''Analysis of vibration isolation systems using a graph model,'' J. Sound Vib. **216**, 399–412 (1998).
[13] W. J. Hsueh, ''Free and forced vibrations of stepped rods and coupled systems,'' J. Sound Vib. **226**, 891–904 (1999).
[14] O. B. Wilson, *Introduction to Theory and Design of Sonar Transducers* (Peninsula, CA, 1988).
[15] S. J. Mason, ''Feedback theory—Further properties of signal flow graphs,'' Proc. IRE **44**, 920–926 (1956).
[16] R. W. Martin and R. A. Sigelmann, ''Force and electrical Thevenin equivalent circuits and simulations for thickness mode piezoelectric transducers,'' J. Acoust. Soc. Am. **58**, 475–489 (1975).

A computational acoustic field reconstruction process based on an indirect boundary element formulation

Zhidong Zhang and Nickolas Vlahopoulos
Department of Naval Architecture and Marine Engineering, The University of Michigan, 2600 Draper Road, Ann Arbor, Michigan 48109-2145

S. T. Raveendra
Automated Analysis Corporation, 2805 South Industrial, Suite 100, Ann Arbor, Michigan 48104-6767

T. Allen
Scientific Research Laboratories, Ford Motor Company, 20000 Rotunda Drive, Building SRL MD2636, Dearborn, Michigan 48121

K. Y. Zhang
Truck VC CAE Department, Advanced Vehicle Technology, Ford Motor Company, 20901 Oakwood Boulevard, MD299, P.O. Box 2053, Dearborn, Michigan 48121

(Received 24 March 2000; revised 10 August 2000; accepted 16 August 2000)

The objective of the work presented in this paper is to develop a computational capability based on the indirect boundary element method (IBEM) for evaluating appropriate velocity boundary conditions on an assembly of piston type sources in order to recreate a prescribed acoustic field. Information for the acoustic pressure of the original acoustic field at certain field points constitutes the input to the developed process. The velocities on the piston type sources are computed from transfer functions evaluated between the field points where the acoustic pressure of the original field is prescribed and the velocity boundary condition on each element of the generic source. The IBEM is employed for computing the transfer functions in order to accommodate the presence of openings in the model and radiation from both sides of the piston type sources. Simulating the generic source as a thin surface that radiates from both sides eliminates the presence of irregular frequencies in the analysis. A singular value decomposition (SVD) solver is integrated with the IBEM computations in order to evaluate the velocity boundary conditions from the transfer functions. The number of field points where the acoustic pressure is defined can be considerably smaller than the number of elements where the velocity is computed. The solution that demonstrates the smallest magnitude is selected from all possible solutions. An algorithm is also developed for identifying the optimum field points where the acoustic pressure of the original field must be prescribed. The optimum field points are selected from a set of prescribed candidate points. The number of optimum points is considered smaller than the number of elements where the velocity is computed. The properties of the transformation matrix and the quality of the reconstruction depend on the location of the field points. Thus, the selection of the optimum points is based on achieving the highest possible orthogonality among the vectors that comprise the range of the transformation matrix. Several validation and application cases are presented.
[S0001-4966(00)04611-7]

PACS numbers: 43.40.Rj [CBB]

I. INTRODUCTION

Several developments have been presented in the past in the area of reconstructing a noise source from acoustic field data. A fast Fourier transform algorithm has been utilized in reconstructing the acoustic field created by sources with simple geometry (planar or cylindrical).[1–3] Green's functions that satisfy a homogeneous condition on the source surface must be available for this development. Thus, the process is applicable only to sources with simple geometry.[1–3] A technique based on the direct boundary element method for identifying a noise source was presented, and the effect of the presence of random error in the field measurements was investigated.[4] Other acoustic holography techniques also based on the direct boundary element method were developed for reconstructing sources with irregular geometry.[5–8] A singular value decomposition (SVD) algorithm was utilized for evaluating velocity on the surface of the vibrating object from the acoustic field data.[5–8] The concept of integrating the direct boundary element method with a SVD algorithm for a source reconstruction was utilized for both exterior radiation[5,6,8] and interior acoustics.[7] Results for sound sources such as a pulsating sphere, a vibrating piston on a rigid sphere, a pulsating cylinder with spherical end caps, and a flexible panel in a baffle have been used for validation.[6,8] Other efforts in the area of acoustic field reconstruction were associated with the development of a Helmholtz equation least squares (HELS) method for the reconstruction of the surface pressure field on sources of irregular

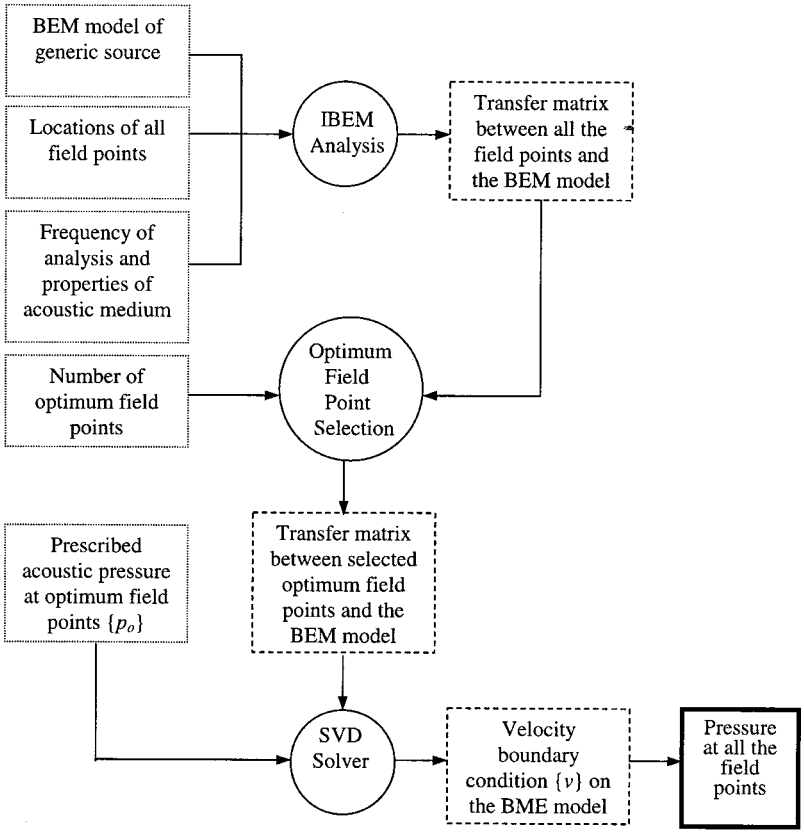


FIG. 1. Flow chart of the field reconstruction process.

shapes for both interior and exterior fields.[9,10] The HELS method is based on an expansion of spheroidal functions, and allows recovery of the pressure field on the surface of the noise sources.

In order to identify the best field points for collecting acoustic data for a source reconstruction computation the effective independence (EFI) method[11,12] was combined with the SVD algorithm and the direct boundary element method.[7,8] The optimal field point selection process identifies linearly independent measurement locations for source reconstruction based on measurements for both an interior[7] and an exterior acoustic field.[8] The EFI method bases the selection of points on the independence of the Fisher information matrix associated with an already orthogonal basis. Consequently the number of selected points has to be higher or equal to the number of vectors in the basis. Therefore, when the EFI was combined with the direct boundary element method for identifying the optimal field points, a requirement was imposed to the number of selected optimal field points to be higher than the number of elements where the velocity was computed.[7,8]

The work presented in this paper targets the computation of appropriate velocity boundary conditions on the surface of a generic source that has approximately the same dimensions with the actual source, in order to recreate the acoustic field that was originally generated by the actual source. Based on the IBEM formulation, transfer functions are defined between the acoustic response at field points, and the velocity boundary conditions on the generic source. A system of equations and a transformation matrix can be assembled from the transfer functions. The velocity boundary conditions on the generic source constitute the unknown variables. A SVD algorithm[13] is employed for solving the final system of equations formulated by the IBEM. An iterative algorithm that selects the optimum location of a prescribed number of points from a set of candidate points is developed. The capability to select the optimum points from a candidate set can be utilized for reducing the number of field points required in the field reconstruction computation. The number of optimum points can be considerably less than the number of elements where the velocity is computed. The flow chart of the developed field reconstruction process is demonstrated in Fig. 1. Several field reconstructions are performed in order to validate the development. The impact on the quality of the results of selecting field points in an optimal manner is demonstrated.

The new developments presented in this paper are: (i) A field reconstruction algorithm based on the IBEM (the IBEM is a substantially different boundary element formulation than the direct boundary element method that has been employed in previous source reconstruction developments). (ii) Development of an algorithm for the selection of optimal field points based on maximizing the quality of the reconstruction by optimizing the orthogonality property of the vectors that comprise the range of the transformation matrix. The number of optimal points is considered to be less than the number of elements where the velocity is computed.

II. MATHEMATICAL FORMULATION

A brief overview of the IBEM is presented first. A SVD algorithm employed for computing the element velocities that recreate the prescribed acoustic field is introduced. A

selection algorithm that identifies the optimum location of a prescribed number of field points from a set of candidate points is also presented.

A. Overview of indirect boundary element method

In the IBEM the primary variables of the formulation are defined as the difference in the pressure, and the difference in the normal gradient of the pressure between the two sides of a boundary element model.[14] The acoustic pressure at any data recovery point can be computed by the surface integral:

$$p(\mathbf{r})=\int_{S_Y}\left(G(\mathbf{r},\mathbf{r}_Y)\,\delta dp(\mathbf{r}_Y)-\frac{\partial G(\mathbf{r},\mathbf{r}_Y)}{\partial \hat{n}_Y}\,\delta p(\mathbf{r}_Y)\right)dS_Y, \tag{1}$$

where S_Y=surface of the boundary element model, subscript ''Y'' indicates a source point on the boundary element surface, $\mathbf{r}$=position vector for the field point, $p(\mathbf{r})$=acoustic pressure at a point defined by position vector $\mathbf{r}$, $\mathbf{r}_Y$=position vector of a source point on the surface of the model, $\hat{n}_Y$=unit normal at the location of the source point, the Green's function, $G(\mathbf{r},\mathbf{r}_Y)$ is defined as $G(\mathbf{r},\mathbf{r}_Y)=(1/4\pi|\mathbf{r}-\mathbf{r}_Y|)\times e^{-j|\mathbf{r}-\mathbf{r}_Y|}$, $\delta dp(\mathbf{r}_Y)$, $\delta p(\mathbf{r}_Y)$=primary variables of the IBEM defined as the difference in the normal gradient of the pressure and the difference in the pressure, respectively, between the two sides of the boundary element model.[14] The difference in the normal gradient of the pressure can be defined

$$\begin{aligned}\delta dp(\mathbf{r}_Y)&=\frac{\partial p(\mathbf{r}_{Y_1})}{\partial \hat{n}_Y}-\frac{\partial p(\mathbf{r}_{Y_2})}{\partial \hat{n}_Y}\\&=-j\rho\omega(v(\mathbf{r}_{Y_1})+v(\mathbf{r}_{Y_2})),\end{aligned} \tag{2}$$

where subscripts ''1'' and ''2'' are associated with the two sides of the boundary element model, ρ=density of the acoustic medium, ω=radial frequency, $v(\mathbf{r}_{Y_1})$=acoustic velocity on side ''1'' of the model at location $\mathbf{r}_Y$, and $v(\mathbf{r}_{Y2})$=acoustic velocity on side ''2'' of the model at location $\mathbf{r}_Y$. The difference in the acoustic pressure across the surface of the boundary element model can be defined

$$\delta p(\mathbf{r}_Y)=(p(\mathbf{r}_{Y_1})-p(\mathbf{r}_{Y_2})). \tag{3}$$

The definition of the primary variables contains information for the acoustic spaces from both sides of the model. Thus, the IBEM can model a thin source that radiates from both sides and can account for openings in the acoustic model. The definition of the primary variables and the assembly of the primary system of equations constitute major differences between the direct and indirect methods.

By considering for simplicity a structural velocity boundary condition over the entire boundary element model (general boundary conditions such as pressure, velocity and unequal impedance are possible[14–16]) the boundary condition becomes

$$v(\mathbf{r}_{X_1})=-v(\mathbf{r}_{X_2})=\overline{v}(\mathbf{r}_X)\Rightarrow\delta dp(\mathbf{r}_X)=0 \quad \text{on} \quad S_v\equiv S_Y, \tag{4}$$

where $\overline{v}(\mathbf{r}_X)$=structural velocity at a location of the boundary element model defined by the point vector $\mathbf{r}_X$, S_Y=surface of the entire boundary element model, and S_v=surface with velocity boundary condition. An integral form of the boundary condition can be created by differentiating Eq. (1), taking into account the relationship between the normal gradient of the pressure and the acoustic velocity, and combining them with Eq. (4),

$$\begin{aligned}-j\rho\omega\overline{v}(\mathbf{r}_X)&=\frac{\partial p(\mathbf{r}_X)}{\partial \hat{n}_X}\\&=\int_{S_Y}\left(-\frac{\partial^2 G(\mathbf{r}_X,\mathbf{r}_Y)}{\partial \hat{n}_X\partial \hat{n}_Y}\,\delta p(\mathbf{r}_Y)\right)dS_Y.\end{aligned} \tag{5}$$

A numerical solution is obtained through the variational principle[17] stating that the solution to the equation,

$$f=\Re(\phi), \tag{6}$$

will also minimize the functional[17]

$$F(\phi)=\int_S \phi\Re(\phi)dS-2\int_S \phi f dS, \tag{7}$$

where $\Re(\phi)$=the operator on function ϕ, f=a known function, and $F(\phi)$=the quadratic functional. By identifying the correspondence between the terms of Eq. (5) and the terms of Eq. (6) a functional can be generated according to Eq. (7),

$$\begin{aligned}F_v=2\int_{S_X} j\rho\omega\overline{v}(\mathbf{r}_X)\,\delta p(\mathbf{r}_X)dS_X-\int_{S_X}\delta p(\mathbf{r}_X)\\\times\left(\int_{S_Y}\left(\frac{\partial^2 G(\mathbf{r}_X,\mathbf{r}_Y)}{\partial \hat{n}_X\partial \hat{n}_Y}\,\delta p(\mathbf{r}_Y)\right)dS_Y\right)dS_X.\end{aligned} \tag{8}$$

By imposing a stationarity condition on F_v with respect to the unknown primary variables $\delta p(\mathbf{r}_X)$ results in a linear system of equations that allows computation of $\delta p(\mathbf{r}_Y)$ over the entire boundary element model.

The unknown functions of the primary variables on the surface of the boundary element model are expressed in terms of the unknown nodal values of the primary variables at the nodes of the boundary element model, and the element shape functions as

$$\delta p(\mathbf{r}_X)=\{N_i\}^T\{\delta p_i\}, \tag{9}$$

where $\{\delta p_i\}$=column vectors of nodal values of the primary variables at the nodes of the ith element, and $\{N_i\}$=column vector of the shape functions associated with the ith element. Spatial derivatives of the primary variables can be also expressed in terms of the nodal values of the primary variables and the spatial derivatives of the shape functions. The final numerical system of equations is derived by imposing a stationarity condition on F_v with respect to the unknown primary variables. It leads to a system of equations

$$[A]\{q\}=\{f\}, \tag{10}$$

where $[A]$=the acoustic system matrix, $\{q\}$=vector of unknown primary variables on the surface of the BEM model, and $\{f\}$=vector representing the excitation, derived from the velocity boundary condition. The system matrix $[A]$ is square

and symmetric since it is derived by a variational formulation. The terms of matrix $[A]$ are generated from the numerical formulation and the stationarity condition associated with the second integral of Eq. (8). The terms of the excitation vector $\{f\}$ are generated from the first integral of Eq. (8).

B. Formulation of the inverse analysis

Once the distribution of the primary variables is computed by Eq. (10), the integral defined in Eq. (1) can be employed to compute the acoustic response at any field point. By taking into account the numerical discretization introduced by Eq. (9), the computations associated with recovering the acoustic pressure at several field points can be written in matrix form

$$\overbrace{\{p_o\}}^{M\times 1}=\overbrace{[DR]}^{M\times N}\overbrace{\{q\}}^{N\times 1}, \tag{11}$$

where $\{p_o\}$=vector of acoustic pressure at ''M'' number of field points, $[DR]$=matrix with entries derived from the integrals of Eq. (1), and ''N''=number of nodal primary variables present in the boundary element formulation. According to Eq. (10) $\{q\}$ can be expressed in terms of $[A]$ and $\{f\}$. Then,

$$\overbrace{\{p_o\}}^{M\times 1}=\overbrace{[DR]}^{M\times N}\overbrace{[A]^{-1}}^{N\times N}\overbrace{\{f\}}^{N\times 1}. \tag{12}$$

The vector $\{f\}$ is generated from the integral $2\int_{S_X} j\rho\omega\overline{v}(\mathbf{r}_x)\delta p(\mathbf{r}_x)dS_X$ of Eq. (8). The integral, and therefore the terms of vector $\{f\}$, are linear functions of the velocity boundary conditions $\overline{v}(\mathbf{r}_X)$. Thus, the vector $\{f\}$ can be expressed as a product between a matrix and the vector of velocity boundary conditions,

$$\overbrace{\{f\}}^{N\times 1}=\overbrace{[FV]}^{N\times K}\overbrace{\{\overline{v}\}}^{K\times 1}, \tag{13}$$

where $\{\overline{v}\}$=vector of velocity boundary conditions applied on the elements of the boundary element model, and ''K'' =number of elements in the boundary element model. The entries of $[FV]$ are derived from integrals of the form

$$\int_{S_X} j\rho\omega\{N_i\}\{N_j\}^T dS_X\rightarrow[FV]. \tag{14}$$

Introducing Eq. (13) into Eq. (12) results in

$$\overbrace{\{p_o\}}^{M\times 1}=\underbrace{\overbrace{[DR]}^{M\times N}\overbrace{[A]^{-1}}^{N\times N}\overbrace{[FV]}^{N\times K}}_{\underbrace{[B]}_{M\times K}}\overbrace{\{\overline{v}\}}^{K\times 1}\Rightarrow\overbrace{\{p_o\}}^{M\times 1}=\overbrace{[B]}^{M\times K}\overbrace{\{\overline{v}\}}^{K\times 1}. \tag{15}$$

The transformation matrix $[B]$ establishes a direct relationship between the velocity boundary conditions on the surface of the generic source and the acoustic pressure at the field points. Equation (15) is employed for computing the velocity on the surface of a generic source in order to recreate a prescribed acoustic field. For practical reasons it is beneficial to utilize the least possible number of field points for collecting information for the acoustic field. Therefore, in the work presented in this paper the number of field points ''M'' is always considered to be less than the number of elements ''K'' where velocities are being computed. The SVD algorithm and the optimum point selection algorithm have been developed based on the expectation that $M<K$. Matrix $[B]$ contains the following information:

(i) frequency of analysis and the fluid properties (ρ,c,ω),
(ii) relative position between the field points and the boundary element model (matrix $[DR]$),
(iii) complex geometry of the source (boundary element matrices $[DR]$, $[A]$),
(iv) contribution of each element to the excitation applied on the acoustic system (matrix $[FV]$).

In the field reconstruction process the information about the acoustic response $\{p_o\}$ originating from the actual source is considered to be prescribed. The velocity boundary conditions $\{\overline{v_o}\}$ that can be applied on a generic source are computed by employing a SVD solution algorithm[13] for solving Eq. (15). It is necessary to employ a SVD algorithm because $[B]$ is not a square matrix and therefore $[B]^{-1}$ does not exist. Since in this paper it is considered that $M<K$, the solution $\{\overline{v_o}\}$ to Eq. (14) is not unique because any linear combination of vectors from the null space can be added to the solution. However, out of all the possible solutions the one selected as final is the solution that exhibits the smallest magnitude $|\{\overline{v_o}\}|$.[13] According to SVD algebra the $M\times K$ matrix $[B]$ can be written as[13]

$$\overbrace{[B]}^{M\times K}=\overbrace{[U]}^{M\times K}\overbrace{[W]}^{K\times K}\overbrace{[V]^T}^{K\times K}, \tag{16}$$

where $[U]$=a $M\times K$ column matrix, $[W]$=a $K\times K$ diagonal matrix with positive or zero diagonal entries, and $[V]^T$=the transverse of a $K\times K$ orthogonal matrix $[V]$. The columns of $[U]$ that correspond to nonzero diagonal entries of $[W]$ are expected to constitute an orthogonal set of basis vectors that span the range of matrix $[B]$. The columns of $[V]$ that correspond to the zero diagonal terms of $[W]$ constitute an orthogonal basis for the null space. Thus, vectors can be appropriately designated to belong either to the basis of the range of the transformation or to the null space depending on the value of the diagonal terms of matrix $[W]$. The selection of the solution that exhibits the lowest magnitude is made by appropriately assigning in the null space the vectors associated with small diagonal entries in matrix $[W]$. Due to the expected orthogonality properties the basis of the transformation and the basis of the null space are considered to satisfy the equation

$$\overbrace{[U]^T}^{K\times M}\overbrace{[U]}^{M\times K}=\overbrace{[I]}^{K\times K}=[V]^T[V]=[V][V]^T. \tag{17}$$

Once $[U]$, $[W]$, and $[V]$ have been computed by the SVD then Eq. (15) can be employed to calculate $\{\overline{v_o}\}$,

$$[B]\{\bar{v}_o\}=\{p_o\}\Rightarrow[U][W][V]^T\{\bar{v}_o\}=\{p_o\}$$
$$\Rightarrow[V][W]^+[U]^T[U][W][V]^T\{\bar{v}_o\}$$
$$=[V][W]^+[U]^T\{p_o\}$$
$$\Rightarrow\{\bar{v}_o\}=[V][W]^+[U]^T\{p_o\}, \quad (18)$$

where $[W]^+$ is the Moore-Penrose pseudoinverse of $[W]$.[18] In this manner there is no need to invert the nonsquare matrix $[B]$ because $[B]^{-1}$ does not exist. Instead, the decomposition of $[B]$ into the basis of the range and the basis of the null space are utilized for solving the underdetermined system of equations. An important aspect of the SVD algorithm is associated with the replacement of the small diagonal terms of matrix $[W]$ with zero entries. The replacement process determines whether to consider the corresponding vector as a component of the basis associated with either the range $[U]$ or the null space $[V]$ of the transformation matrix $[B]$. Vectors of the range $[U]$ correspond to nonzero diagonal entries and vectors of the null space correspond to zero diagonal entries. The diagonal terms of $[W]^+$ that correspond to the zero entries of $[W]$ are also replaced by a zero entry. This last substitution results in obtaining, out of all possible solutions, the one with the smallest length when $M<K$.[13]

C. Selection of optimum field points for the field reconstruction

In previous work the EFI method[11,12] was combined with the direct BEM for selecting the best field points of an interior[7] or an exterior[8] acoustic field in order to collect data for reconstructing the source. The EFI method was originally developed for placing sensors on space structures for modal identification on orbit.[11,12] The EFI formulation originates from requiring the modal degrees of freedom computed by a finite element analysis to be equal to the modal degrees of freedom identified by the sensors. The number of sensors is always considered to be at the very least, equal to the number of modal degrees of freedom that will be retrieved. Thus, the EFI algorithm deletes redundant information by eliminating sensor locations that do not contribute significantly to the independent information contained within the target mode partitions. The sensors that are eliminated are identified by generating the Fisher information matrix from the already orthogonal modal matrix. The absolute identification space is computed by evaluating the eigenvalues and eigenvectors of the Fisher information matrix. The original orthogonal modal matrix and the basis vectors of the absolute identification space are utilized for calculating the fractional contribution of each sensor location to each eigenvector of the absolute identification space. Based on the fractional contribution of each sensor it is determined which sensors offer the highest independent information to the already orthogonal modal matrix. Since the EFI formulation is structured to identify independent information from an already orthogonal modal matrix, it cannot accommodate identification of a smaller number of optimum sensor locations than the number of modes contained in the original orthogonal matrix. The implementation of the EFI method to the source reconstruction processes for interior[7] or exterior[8] acoustic fields imposes a similar requirement on identifying a larger number of optimum field points than number of elements where the velocity is reconstructed.

In this work an alternative iterative formulation is developed in order to identify optimal field points for collecting acoustic pressure data. The number of optimal field points is considered smaller than the number of elements where velocities are computed. Instead of using the independence of the information in an already orthogonal matrix as a selection criterion, the developed algorithm is based on selecting field points that maximize the quality of the reconstruction. The fidelity of the reconstructed field is high when the orthogonality demonstrated by the range of the transformation matrix is high. The location of the field points impacts the selection of the vectors that constitute the basis of the range of the transformation matrix. Therefore, achieving the highest possible orthogonality is utilized as criterion for selecting the optimal field points. The new algorithm can be utilized when the number of the optimum field points is lower than the number of elements where the velocity is computed.

The starting point for the optimum field point selection algorithm is the requirement for the reconstructed field pressure to be the same with the baseline field pressure. The acoustic pressure for the recreated field can be evaluated by combining Eq. (15) and Eq. (18),

$$\{p_r\}=[B]\{\bar{v}_o\}=[B][V][W]^+[U]^T\{p_o\}$$
$$=[U][W]\underbrace{[V]^T[V]}_{[I]}[W]^+[U]^T\{p_o\}$$
$$=\underbrace{[U]}_{M\times K}\underbrace{[W]}_{K\times K}\underbrace{[W]^+}_{K\times K}\underbrace{[U]^T}_{K\times M}\{p_o\}, \quad (19)$$

where $\{p_r\}$=vector of reconstructed acoustic pressure from the generic source. Since M is less than K, in order to obtain out of all possible selections the one with the smallest length, the small diagonal terms of $[W]$ and the equivalent positions of $[W]^+$ are replaced by zero terms. In this manner the product of $[W]$ and $[W]^+$ is not equal to the identity matrix, but equal to a $K\times K$ matrix that has unit entries at the diagonal locations that correspond to the vectors of the range of the transformation and zero entries everywhere else. Therefore, the product,

$$\underbrace{[T]}_{M\times M}=\underbrace{[U]}_{M\times K}(\underbrace{[W][W]^+}_{K\times K})\underbrace{[U]^T}_{K\times M}, \quad (20)$$

constitutes a transfer matrix between the original and the reconstructed fields. The right-hand side of Eq. (20) depends on the orthogonality property demonstrated by the basis of the range of the transformation. Therefore, when the orthogonality property is high $[T]$ tends to be equal to the identity matrix and the acoustic field is reconstructed accurately. Each diagonal entry of matrix $[T]$ is related to a particular field point. The selection of the location of the field points influences the derivation of the range and the orthogonality property demonstrated by its basis vectors. An iterative process is established for selecting the optimum location for a prescribed number of field points from a set of candidate points in order to maximize the orthogonality property

of the range of the transformation. Within each iteration matrix $[T]$ is computed, and the field point that corresponds to the lowest diagonal value is eliminated from the set of candidate points. For the remaining points matrices $[U]$, $[W]$, and $[V]$ are evaluated again, and the product $[T]$ is computed. With each iteration the criterion of the smallest resulting velocity vector is utilized in order to determine whether a vector will be assigned to the basis of the range or the basis of the null space of the transformation matrix. The elimination process continues until the number of the remaining field points is equal to the number of points requested to be identified by the optimum search process. The developed optimum search process can also be utilized to identify the best field points over a frequency range. An extra inner loop over the frequency range is added to the iteration process. The elimination of the field points is based on the lowest sum of magnitudes of the diagonal terms over a frequency range rather than at a single frequency. The identification of the optimum locations for the field points is strictly a function of the geometry of the generic source, the frequency of analysis, the acoustic medium, and the relative position of the candidate field points with respect to the source. The selection of optimum field points does not depend on the actual pressure values of the original field, since $\{p_o\}$ does not enter into the computations associated with the derivation of $[U]$, $[W]$, and $[V]$.

III. VALIDATION / APPLICATION

The field reconstruction capability presented in this paper and the optimum field point selection algorithm are validated extensively. Four different types of acoustic fields derived from analytical solutions are utilized in the validation for creating the baseline field information. Velocity boundary conditions are computed on generic sources in order to recreate the baseline fields. A pulsating line source of finite length and a vibrating piston set in a rigid sphere are two examples utilized in the past for validating a source reconstruction development[6] and they are also employed in this paper. In addition, the analytical solution of a quadrupole and the analytical solution of an array of point sources positioned in a plane are utilized for demonstrating the new computational capabilities. The unique capability of the IBEM to capture radiation from both sides of a boundary element model allows definition of thin planar surfaces as generic sources. Appropriate velocity boundary conditions are computed through the developed process. For all four types of acoustic fields a considerably smaller number of field points than number of elements in the boundary element model are utilized in the reconstruction. The field points are identified by the optimal field point selection algorithm presented in this paper. Results for the reconstructed field based on the optimally selected points are compared to the baseline solution and to reconstruction results based on an equal number of evenly distributed field points. The baseline and the reconstructed fields are computed at a larger number of field points than the ones actually used in the reconstruction. Therefore, the correlation is inspected for the overall radiated field and at points that do not participate in the reconstruction process. A metric based on the radiated acoustic power is used in order to compare the solutions. In addition, the radiation patterns and the sound pressure levels are also compared during the validation.

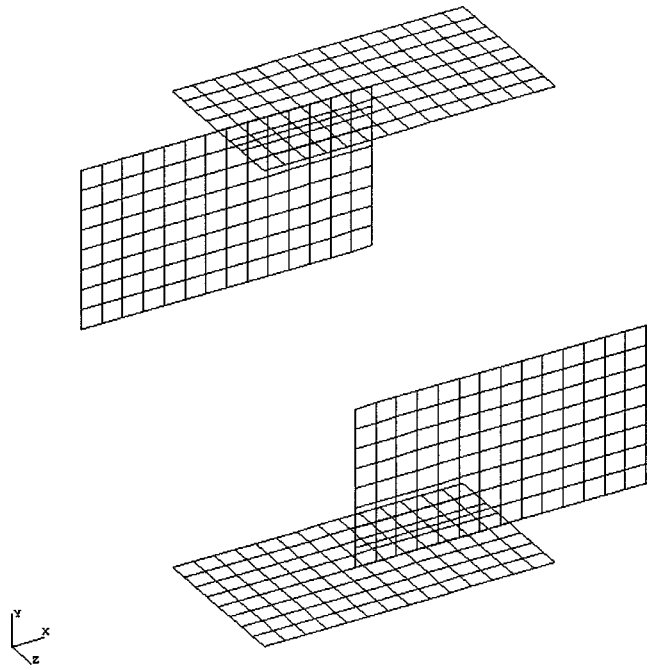


FIG. 2. Boundary element model of the generic source utilized to reconstruct the field generated by a line source.

A. Reconstruction of acoustic field originating from line source

The acoustic field generated by a line source of finite length is considered in this analysis. The line source has length equal to 0.5 m, radius equal to 0.01 m, and radial velocity equal to 100 m/sec. The parameters for the line source are identical to the values utilized in Ref. 6. In order to demonstrate the unique capabilities that the IBEM provides in the field reconstruction process the generic source is considered to be composed of four planes with dimensions 0.6×1.1 m each. Each plane comprises 112 elements that radiate noise from both sides. The four planes are positioned at 0.6 m apart and the boundary element model is depicted in Fig. 2. Since the generic source does not contain an enclosed acoustic cavity there are no irregular frequencies present in the analysis. The capability to model thin sources that radiate noise from both sides is unique to the IBEM and to the field reconstruction technique presented in this paper. The field surface is a sphere with radius equal to 3.0 m and center at the middle of the line source. It is comprised of 98 field points. Reconstruction analyses are performed at $ka=2$ and $ka=4$ similar to Ref. 6 as well as $ka=0.5$, $ka=1$, and $ka=3$. Thirty field points are selected optimally by the developed algorithm for $ka=0.5$, $ka=1$, and $ka=2$ while 50 field points are selected optimally for $ka=3$ and $ka=4$. Values of the original acoustic field at the optimally selected field points are employed to reconstruct the velocity boundary conditions on the generic source. The acoustic field is reconstructed and results are computed at all 98 field points. In order to demonstrate the importance of selecting the field points in an optimal manner the field reconstruction compu-

TABLE I. Radiated acoustic power for continuous line source.

ka	Original (watts)	Optimal (watts)	Error (%)	Even (watts)	Error (%)
0.5	924.75	920.36	0.47	891.34	3.61
1	3069.7	3003.6	2.15	3356.2	9.33
2	6519.7	6378.5	2.17	7008.0	7.49
3	9394.4	9241.5	1.63	8003.5	14.8
4	14 490	14 326	1.13	11 683	19.4

tations are repeated based on acoustic data collected at an equal number of evenly distributed field points. The radiated acoustic power is computed from acoustic pressure and acoustic velocity data calculated at all 98 field points and the results are summarized in Table I. It can be observed that when the optimally selected points are utilized in the computations the maximum error encountered in the radiated power is 2.17%, while for an equal amount of evenly distributed points the maximum error is 19.4%. These results demonstrate that achieving the highest possible orthogonality in the basis of the range of the transformation matrix impacts significantly the quality of the reconstructed field. The developed algorithm for the optimal point selection is based on maximizing the orthogonality property. Therefore, there is significant difference in the results between optimally selected and evenly selected field points.

B. Reconstruction of acoustic field originating from a vibrating piston set in a rigid sphere

A sophisticated source utilized by Ref. 6 is also employed in this paper to further validate and demonstrate the new developments. A vibrating piston is set in the side of a rigid sphere with radius equal to 0.2 m. The spherical piston surface spans an angle of 30°. The piston velocity is defined equal to 1 m/sec. The field pattern and the directivity depend on the Helmholtz number ka. Analyses are performed at $ka=2$ and $ka=4$. The properties of the source and the frequencies of analysis are identical with the values utilized in Ref. 6. The generic source is comprised in this case by a cube with dimension equal to 0.4 m. The center of the cube is the same with the center of the sphere. The boundary element model comprises 150 elements and 152 nodes. A spherical field surface composed of 98 points is defined with

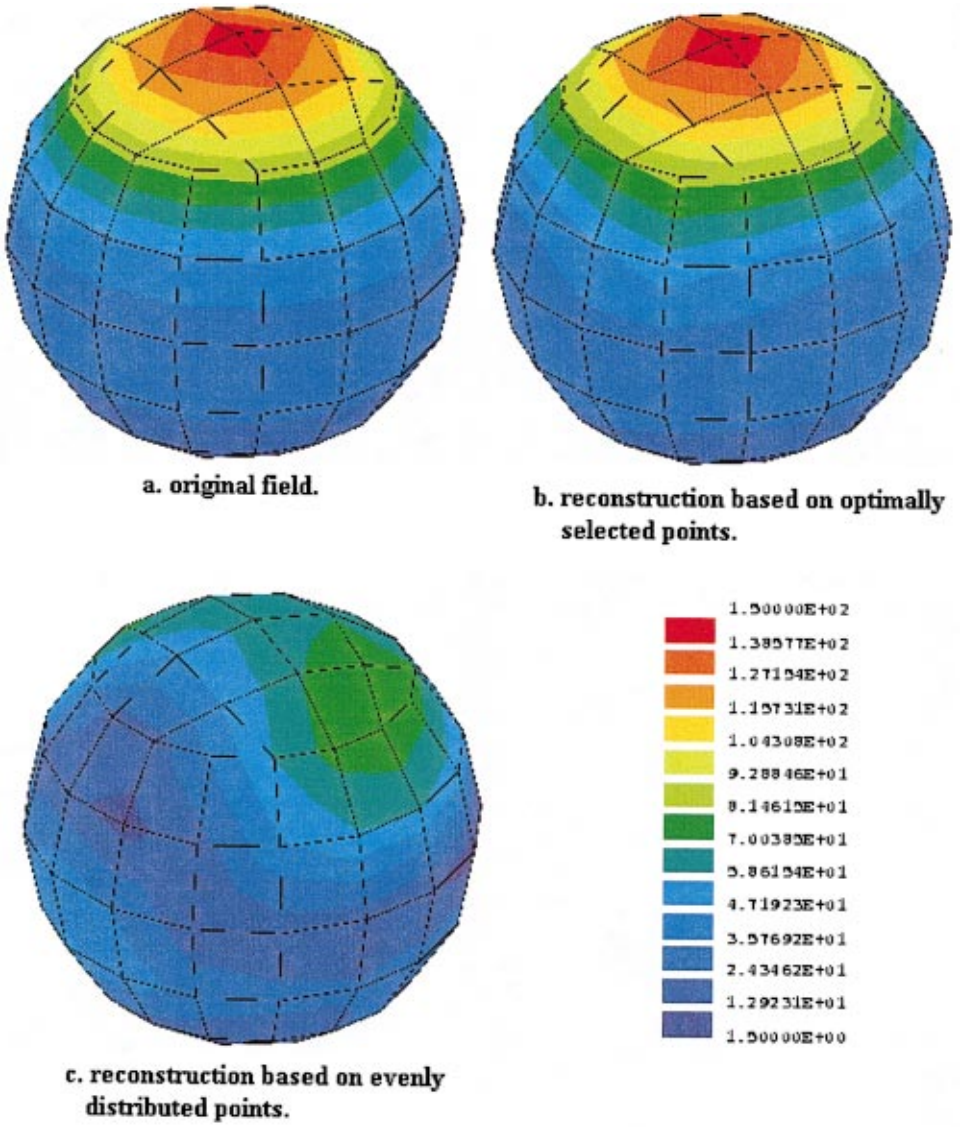


FIG. 3. Magnitude of acoustic pressure of original and reconstructed field; vibrating piston in a sphere; $ka=2$.

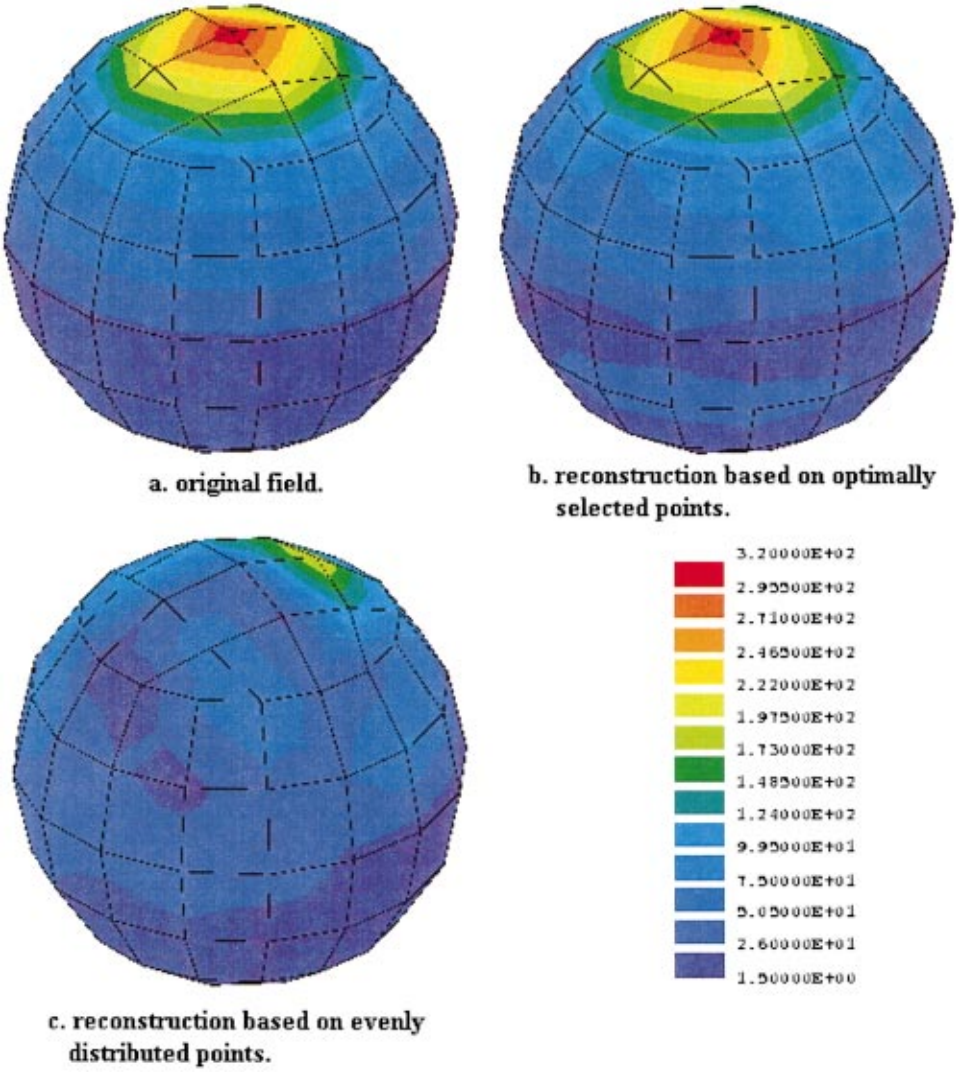


FIG. 4. Magnitude of acoustic pressure of original and reconstructed field; vibrating piston in a sphere; $ka=4$.

0.4 m radius and a center coinciding with the center of the rigid sphere and the cube. The field reconstructions for $ka=1$ are performed based on field data collected at 30 optimally selected and evenly distributed points. The computations $ka=4$ are based on field information collected at 45 field points. In both cases the number of field points is considerably smaller than the number of elements in the boundary element model. The optimal field points are selected from the original 98 points that comprise the field surface. Results for the acoustic pressure at all 98 field points and for Helmholtz numbers $ka=2$ and $ka=4$ are presented in Figs. 3 and 4, respectively. In each figure the magnitude of the acoustic pressure from the baseline field and the two reconstructed fields are presented. The same color scale is employed for all the results in each figure. As expected, the field reconstructed from acoustic data collected at the optimally selected field points produces accurate results for both Helmholtz numbers. The generic source recreates the acoustic field well both at the optimally selected field points that are utilized in the computations, and at field points where the acoustic pressure is not considered during the reconstruction process. The quality of the results computed from acoustic data collected at evenly distributed points is inferior in terms of both the magnitude and the radiation pattern of the reconstructed field. Therefore, the importance of selecting the field points in an optimal manner that achieves the highest possible orthogonality for the range of the transformation matrix is demonstrated.

C. Reconstruction of acoustic field originating from quadrupole

A quadrupole source is utilized to create a baseline acoustic field at 100, 300, 500, 700, and 1000 Hz, respectively. The spacing between two adjacent point sources is 0.01 m. Each one of the four point sources has unit strength and is 180° out of phase with adjacent source. A boundary element model of a box with dimensions $0.2\times0.2\times0.2$ m is comprised of 24 elements, and it is employed as a generic source in the analysis. A spherical field surface of 1 m radius composed of 98 field points is defined. Nine optimally selected and nine evenly distributed points are utilized in two separate reconstruction computations. Figure 5 depicts the two sets of points on the data recovery surface. The magnitude of the acoustic pressure of the baseline acoustic field and the two reconstructed fields based on optimally selected and evenly distributed field points are presented in Fig. 6 for 500 Hz. When the optimally selected field points are em-

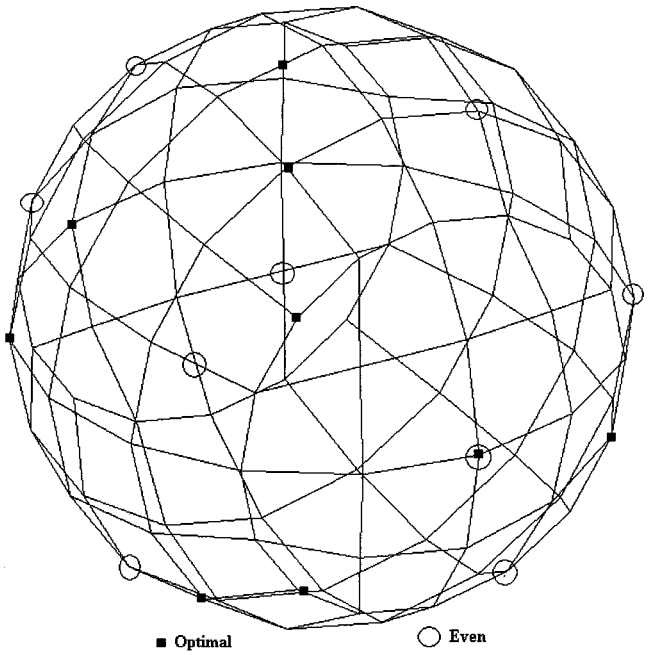


FIG. 5. Distribution of the optimally selected and evenly distributed field points utilized for reconstructing a field created by a quadrupole.

ployed in the reconstruction, the results are identical with the analytical solution for all 98 points (although only 9 points are employed in the reconstruction). For the evenly selected points the results deteriorate. The difference in the reconstructed results originates from the transfer matrix $[T]$ [Eq. (20)] that relates the original and the reconstructed fields. When the range of the transformation matrix $[B]$ comprises an orthogonal set of vectors, then the transfer matrix $[T]$ is expected to be equal to the identity matrix. For the optimally selected points $[T]$ is in fact equal to the identity matrix. For the evenly distributed points the diagonal entries of $[T]$ range from 0.7 to 1.0, and the nonzero off-diagonal terms range between -0.3 and 0.3. The capability to optimally select points is important when it is desired to limit the total number of field points employed in the reconstruction. The situation is similar for the field reconstruction at all other frequencies. Results for the radiated power computed on the spherical field surface are summarized in Table II. Results for the baseline and the reconstructed fields based on the even and optimal distribution of field points are presented. The maximum error encountered in the results from the op-

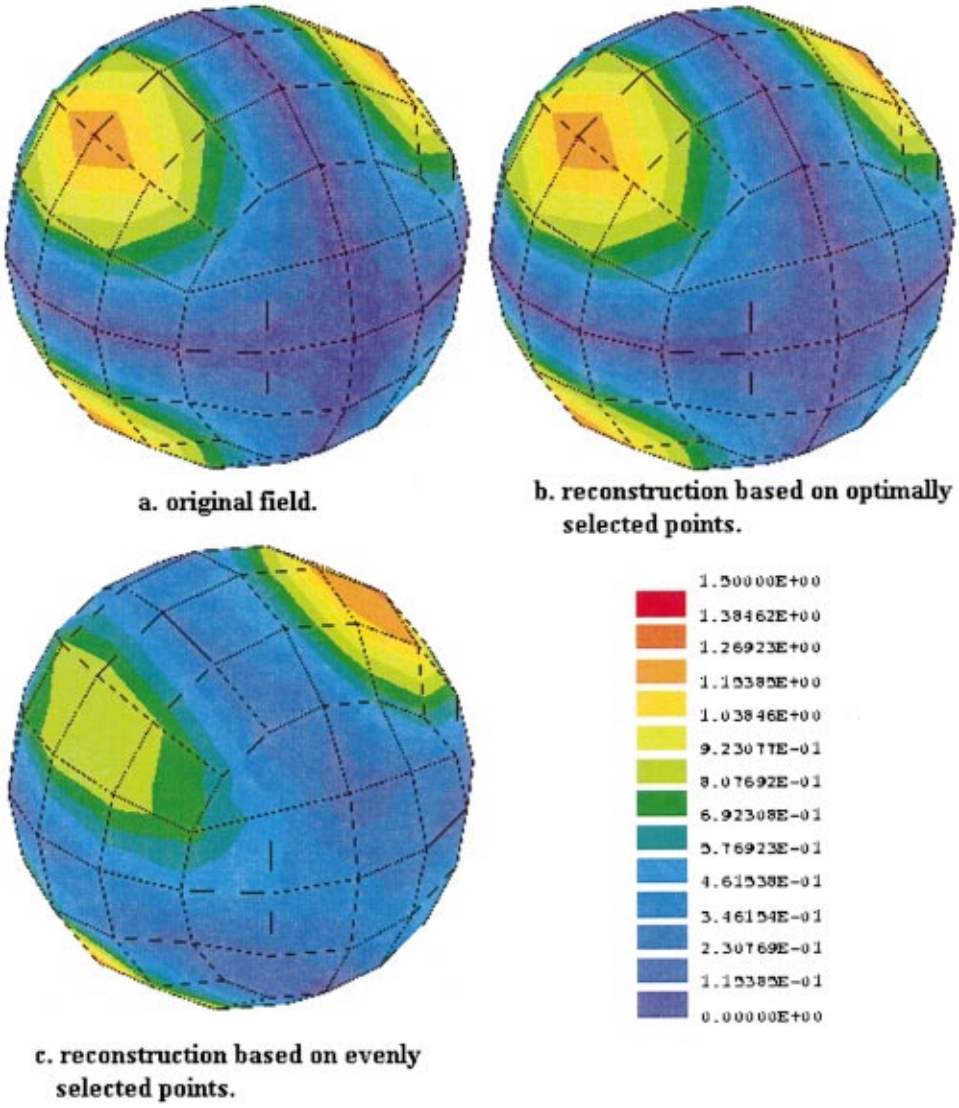


FIG. 6. Magnitude of acoustic pressure of original and reconstructed field; quadrupole field; 500 Hz.

TABLE II. Radiated acoustic power for quadrupole source.

Frequency (Hz)	Original (watts)	Optimal (watts)	Error (%)	Even (watts)	Error (%)
100	3.811E-7	3.816E-7	0.118	3.756E-7	1.44
300	2.782E-4	2.775E-4	0.252	2.933E-4	5.43
500	0.005 962	0.005 993	0.520	0.005 597	6.12
700	0.044 86	0.044 62	0.535	0.045 87	2.25
1000	0.3809	0.3842	0.866	0.3721	2.31

timally distributed points is 0.87%, while the maximum error encountered in the results from the evenly distributed points is 6.12%.

D. Reconstruction of acoustic field originating from a planar distribution of multiple point sources

An acoustic field generated by multiple point sources positioned on a plane is employed in this analysis. The point sources are equally spaced on a plane with dimensions 0.2 by 0.4 m. The spacing between two adjacent point sources is 0.02 m. Each point source has unit strength. The phase of adjacent sources varies by 180°. A spherical field point surface with 0.6 m radius and center positioned at the center of the plane with the sources is defined. The field surface is composed of 98 points. The generic source is defined by two perpendicular planes (Fig. 7). The length of the planes is equal to 0.4 m and the width is equal to 0.2 m. Noise is radiated from both sides of the two planes. Multiple connection and free edge conditions are present in the boundary element model. Field reconstruction analyses are performed at 100, 300, 500, 700, and 1000 Hz, respectively. The reconstructions are based on field data collected at 60 optimally and evenly selected points. All the points of the field surface are considered as candidate points. The optimal points are selected based on a multifrequency criterion. The even and the optimal distributions of field points are presented in Fig. 8. Results for the radiated sound power of the original field and the reconstructed fields are summarized in Table III. The maximum error in the sound power computation for the reconstruction based on optimally distributed points is 4.3%, while the maximum error for the reconstruction based on evenly distributed points is 22%. The sound pressure level

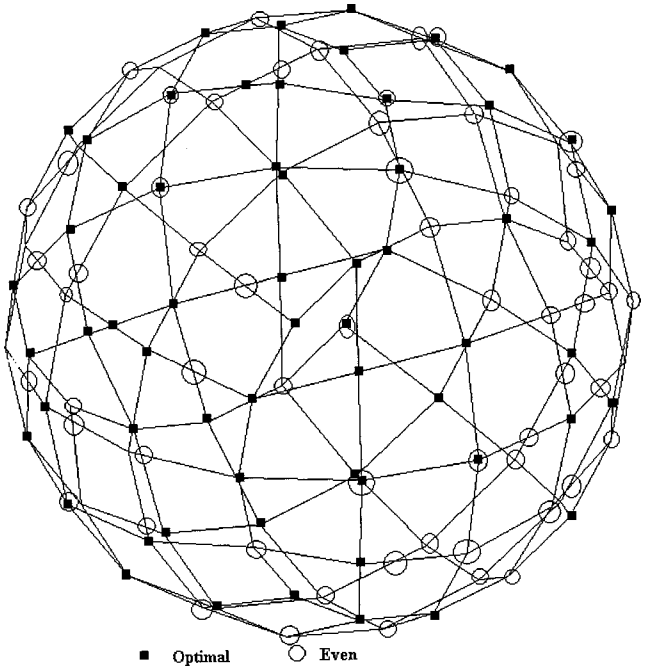


FIG. 8. Evenly and optimally distributed field points utilized to reconstruct the field generated by multiple point sources on a plane.

from the original field and the two reconstructed fields are presented in Fig. 9 for 1000 Hz. Similar to the previous applications it can be observed that the field reconstruction based on the optimally selected points produces considerably more accurate results than the field reconstruction based on an equal amount of evenly distributed points. The results for both the magnitude of the sound power level and the radiation pattern are inaccurate when the evenly distributed field points are employed in the reconstruction. The results improve dramatically when the same number of field points is optimally distributed. In this application the generic source is comprised from thin surfaces that radiate from both sides. Therefore, no irregular frequencies will appear in the analysis and the field reconstruction computations.

IV. CONCLUSIONS

A field reconstruction process is presented in this paper. Velocity boundary conditions are computed on a generic

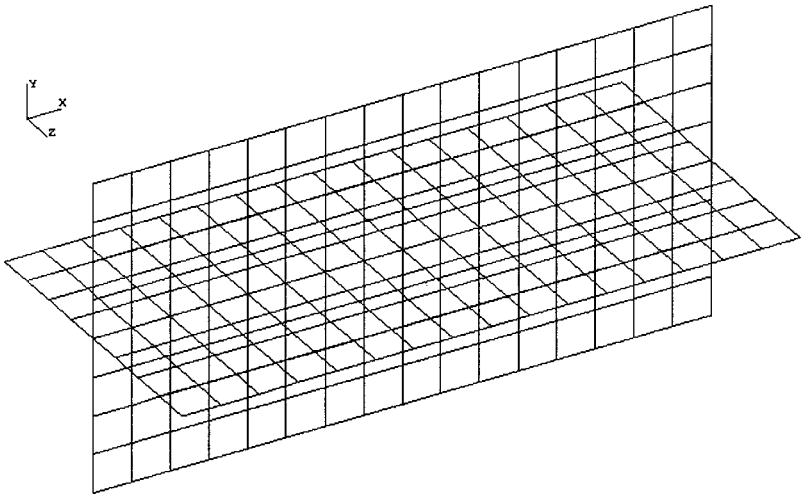


FIG. 7. Boundary element model of generic source utilized to reconstruct the field generated by multiple point sources on a plane.

TABLE III. Radiated acoustic power for multiple point sources in plane.

Frequency (Hz)	Original (watts)	Optimal (watts)	Error (%)	Even (watts)	Error (%)
100	0.1616	0.1553	3.90	0.1334	17.5
300	53.808	56.131	4.32	61.234	13.8
500	700.38	678.21	3.17	592.03	15.5
700	2685.3	2584.5	3.75	2231.4	16.9
1000	6353.7	6153.2	3.16	4930.3	22.4

source in order to recreate a prescribed acoustic field. The development is based on computing transfer functions between velocity boundary conditions on the surface of the generic source and the field points where the baseline field is defined. The computation of the transformation matrix is based on the IBEM. The IBEM offers the significant and unique capabilities of capturing radiation from both sides of a thin surface and including openings and multiple connections in the model. Using a thin open surface as a generic source eliminates the presence of irregular frequencies from the boundary element computations. A SVD algorithm is employed for computing the boundary conditions on the generic source and reconstructing the field. The original field is considered to be prescribed at a limited number of field points that is smaller than the number of elements in the boundary element model. Since an underdetermined system of equations must be solved, from all possible solutions the one that demonstrates the smallest magnitude is selected. It is demonstrated that the quality of the reconstructed field depends on the orthogonality exhibited by the range of the transformation matrix. The locations of the field points where the acoustic pressure of the baseline field is defined determines the orthogonality property of the range of the transformation matrix. An algorithm is developed that identifies the optimal locations of a prescribed number of field points from a set of candidate points. The selection algorithm is based on achieving the highest possible orthogonality for the basis of the range of the transformation matrix. The developments are validated by reconstructing the acoustic field generated from four different types of sources. The radiated sound power, the radiation pattern, and the sound pressure level of the radiated noise constitute metrics for the correlation. Good agreement is observed between the baseline field and the field reconstructed from acoustic data collected at optimally selected field points. The significance of field

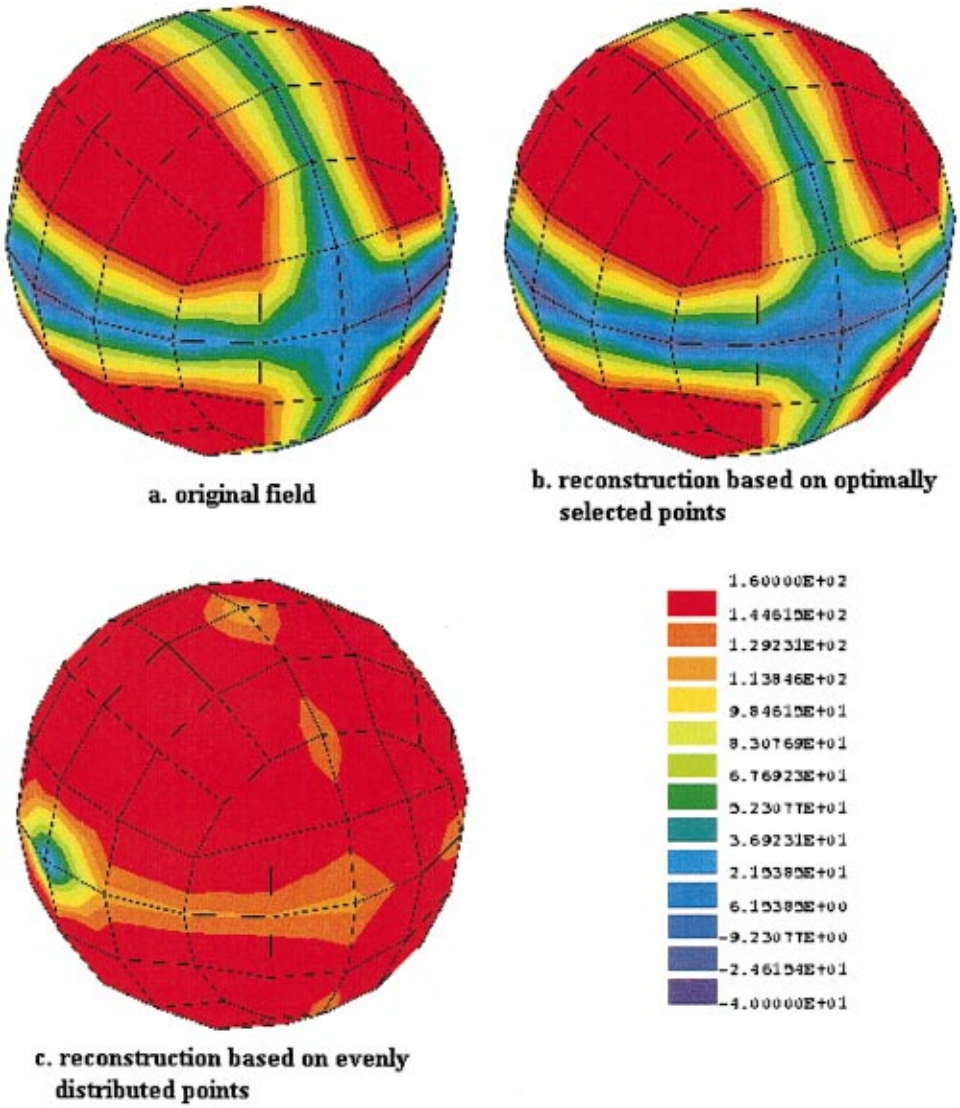


FIG. 9. Sound pressure level of original and reconstructed field; planar distribution of multiple point sources; 1000 Hz.

points selection on the quality of the results is demonstrated by comparing fields reconstructed from acoustic data collected at the same number of evenly and optimally distributed field points. The results computed from acoustic data collected at the optimal field points exhibit consistently high correlation with the baseline field.

ACKNOWLEDGMENTS

The work presented in this paper was funded by the Ford Motor Company University Research Program.

[1] J. D. Maynard, E. G. Williams, and Y. Lee, ''Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH,'' J. Acoust. Soc. Am. **78**, 1395–1413 (1985).
[2] W. A. Veronesi and J. D. Maynard, ''Nearfield acoustic holography (NAH): II. Holographic reconstruction algorithms and computer implementation,'' J. Acoust. Soc. Am. **81**, 1307–1322 (1987).
[3] E. G. Williams and H. D. Dardy, ''Generalized nearfield acoustical holography for cylindrical geometry: Theory and experiment,'' J. Acoust. Soc. Am. **81**, 389–407 (1987).
[4] B. K. Gardner and R. J. Bernhard, ''A noise source identification technique using an inverse Helmholtz integral equation method,'' J. Vib. Acoust. Stress Reliab. Design **110**, 84–90 (1988).
[5] W. A. Veronesi and J. D. Maynard, ''Digital holographic reconstruction of sources with arbitrarily shaped surfaces,'' J. Acoust. Soc. Am. **85**, 588–598 (1989).
[6] M. R. Bai, ''Application of BEM (boundary element method)-based acoustic holography to radiation analysis of sound sources with arbitrarily shaped geometries,'' J. Acoust. Soc. Am. **92**, 533–549 (1992).
[7] B. K. Kim and J. G. Ih, ''On the reconstruction of the vibro-acoustic field over the surface enclosing an interior space using the boundary element method,'' J. Acoust. Soc. Am. **100**, 3003–3015 (1996).
[8] J. G. Ih and B. K. Kim, ''On the use of the BEM-based NAH for the vibro-acoustic source imaging on the nonregular exterior surfaces,'' Proceedings of Noise-Con 98, Ypsilanti, Michigan, pp. 665–770.
[9] S. F. Wu and J. Yu, ''Reconstructing interior acoustic pressure field via Helmholtz equation least-squares method,'' J. Acoust. Soc. Am. **104**, 2054–2060 (1998).
[10] S. F. Wu and Y. Wu, ''Reconstruction of radiated acoustic pressure fields from a complex vibrating structure,'' Proceedings of the ASME Noise Control and Acoustics Division, ASME, 1998, Vol. 25, pp. 339–345.
[11] D. C. Kammer, ''Sensor placement for on-orbit model identification and correlation of large space structures,'' J. Guid. Control Dyn. **14**, 251–259 (1991).
[12] D. C. Kammer, ''Effect of model error on sensor placement for on-orbit modal identification of large space structures,'' J. Guid. Control Dyn. **15**, 334–341 (1992).
[13] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN—The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1987).
[14] N. Vlahopoulos and S. T. Raveerdra, ''Formulation, implementation, and validation of multiple connection and free edge constraints in an indirect boundary element formulation,'' J. Sound Vib. **210**, No. 1, 137–152 (1998).
[15] J. P. Coyette and K. R. Fyfe, ''Solution of elasto-acoustic problems using a variational finite element/boundary element technique,'' Numerical Techniques in Acoustic Radiation, Proceedings of Winter Annual Meeting of the ASME, San Francisco, Dec. 1989, pp. 15–25.
[16] S. T. Raveendra, N. Vlahopoulos, and A. Glaves, ''An indirect boundary element formulation for multi-valued impedance simulation in structural acoustics,'' Appl. Math. Model. **22**, 379–393 (1998).
[17] S. G. Mikhlin, *Variational Methods in Mathematical Physics* (MacMillan, New York, 1964).
[18] G. Strang, *Linear Algebra and its Applications*, 2nd ed. (Academic, New York, 1980).

Dispersion and characteristic surfaces of waves in laminated composite circular cylindrical shells

Z. C. Xi, G. R. Liu, K. Y. Lam, and H. M. Shang
Department of Mechanical and Production Engineering, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260

(Received 11 April 2000; revised 11 August 2000; accepted 18 August 2000)

The dispersion behaviors and characteristic surfaces of waves in a laminated composite circular cylindrical shell are investigated using a semianalytical method based on the theory of three-dimensional elasticity. The radial displacement of the shell is modeled by finite elements, while the axial and circumferential displacements are expanded as the complex exponentials. The associated characteristic equation is developed by means of the Hamilton's principle. The eigenvalues are established in terms of the Rayleigh quotient. Six characteristic wave surfaces, viz., the phase velocity, phase slowness, and phase wave surfaces, as well as the group velocity, group slowness, and group wave surfaces, are introduced to visualize the effects of anisotropy on wave propagation. Numerical examples demonstrate that the ratio of the inner radius to the thickness of the shell has a stronger influence on the frequency spectra in the circumferential wave than on that in the axial wave; that negative group velocity appears at a range of smaller wave numbers and the range varies as the wave normal and the ratio of the inner radius to the thickness of the shell; and that the characteristic wave surfaces vary with the propagation modes of waves, the ratio of the inner radius to the thickness of the shell, and the lay-ups of the laminated shells. © *2000 Acoustical Society of America.* [S0001-4966(00)03911-4]

PACS numbers: 43.40.Ey, 43.20.Tb [CBB]

I. INTRODUCTION

Wave propagation in anisotropic media has been a subject of intense research interest. Rayleigh, Lamb, Love, and Stoneley waves in anisotropic media have been well studied (Mal, 1988; Datta *et al.*, 1988a, 1988b; Kline, 1992; Nayfeh, 1995; Rose, 1999). Because of the algebraic complexity involved in shells, however, studies of waves in anisotropic shells are still rare. Markus and Mead (1995a, 1995b), and Yuan and Hsieh (1998) studied analytically free waves in composite circular cylindrical shells. Xi *et al.* (1996, 1997a, 1997b, 1999) treated semianalytically free waves of laminated composite shells of revolution either in vacuum or partially filled with a fluid. In their analyses, the effects of transverse shear deformation, material nonlinearity, the coupling between symmetric and antisymmetric modes, and the coupling between the fluid and shell were taken into account. Nelson *et al.* (1971) proposed a numerical-analytical method (NAM) for analyzing waves in laminated orthotropic cylinders. In the analysis, the circumferential and axial displacements are represented by trigonometric functions, while the radial displacement is modeled by finite elements. Huang and Dong (1984) later applied the NAM to study the frequency dispersion of laminated composite cylinders. After that, the NAM proved expedient in the treatment of vibration and wave motion problems of anisotropic cylindrical shells. Rattanwangcharoen *et al.* (1994) attacked the reflection problem of waves at the free edge of laminated circular cylinders. Rattanwangcharoen *et al.* (1997) and Zhuang *et al.* (1997) pursued axisymmetric guided wave scattering by cracks in welded steel pipes. Xi *et al.* (2000a, 2000b, 2000c) examined waves scattered by a crack in fluid-loaded axisymmetric laminated composite cylinders. Han *et al.* (1999) dealt with transient responses of multilayered cylindrical shells made of functionally gradient materials.

This paper attempts to investigate such unique and important wave properties of anisotropic cylindrical shells as group velocity and characteristic wave surfaces using a similar semianalytical method to the NAM proposed by Nelson (1971). The present method is formulated within the framework of the theory of three-dimensional elasticity. The radial displacement of the shell is modeled by finite elements, while the axial and circumferential displacements are expanded as the complex exponentials. The associated dispersion equation is developed by way of the Hamilton's principle. The body forces of the cylindrical shell are neglected. The eigenvalues are established in terms of the Rayleigh quotient. Six representation wave surfaces viz., the phase velocity, phase slowness, and phase wave surfaces as well as the group velocity, group slowness, and group wave surfaces are introduced to visualize the effect of anisotropy on wave propagation. Parametric studies include the propagation direction and modes of waves, the ratio of the inner radius to the thickness of the shell, and the lay-ups of the laminated shell.

II. DISPERSION EQUATION

Suppose that a helical wave propagates in a laminated composite circular cylindrical shell of inner radius R_i and outer radius R_o, where β is the angle of the propagation direction of the wave with respect to the z-axis (see Fig. 1). The perfect bonding between plies is assumed. Deformations

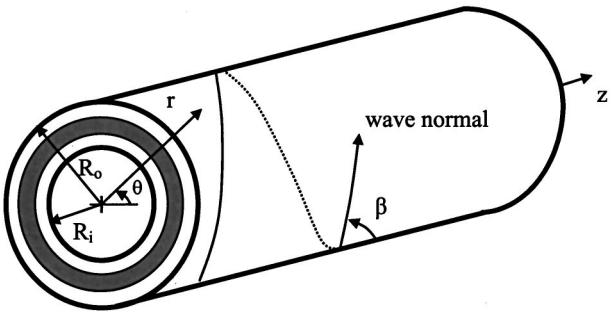


FIG. 1. Configuration of a laminated composite circular cylindrical shell.

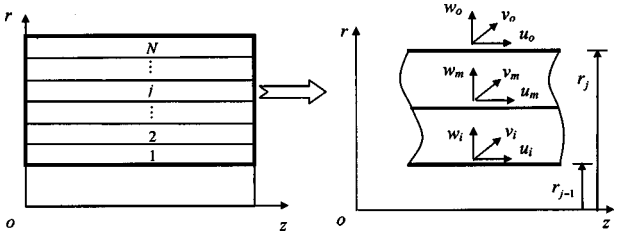


FIG. 2. Annular element subdivision and the *j*th isolated annular element.

of the shell are assumed to be small. Under these assumptions, the strain-displacement relations in the cylindrical coordinate system are given by

$$\boldsymbol{\varepsilon}=\mathbf{L}\mathbf{u}, \tag{1}$$

where $\varepsilon=[\varepsilon_z\varepsilon_\theta\varepsilon_r\gamma_{r\theta}\gamma_{rz}\gamma_{z\theta}]^T$ is the vector of strains and $\mathbf{u}=[uvw]^T$ is the vector of displacements. Here u, v, and w are the displacement components in the axial, circumferential, and radial directions, respectively. The differential operator matrix $\mathbf{L}$ is given by

$$\mathbf{L}=\begin{bmatrix} \frac{\partial}{\partial z} & 0 & 0 & 0 & \frac{\partial}{\partial r} & \frac{1}{r}\frac{\partial}{\partial\theta} \\ 0 & \frac{1}{r}\frac{\partial}{\partial\theta} & 0 & \frac{\partial}{\partial r}-\frac{1}{r} & 0 & \frac{\partial}{\partial z} \\ 0 & \frac{1}{r} & \frac{\partial}{\partial r} & \frac{1}{r}\frac{\partial}{\partial\theta} & \frac{\partial}{\partial z} & 0 \end{bmatrix}^T$$

$$=\mathbf{L}_1\frac{\partial}{\partial z}+\mathbf{L}_1\frac{1}{r}\frac{\partial}{\partial\theta}+\mathbf{L}_3\frac{\partial}{\partial r}+\mathbf{L}_4\frac{1}{r}, \tag{2}$$

where the matrices $\mathbf{L}_1$, $\mathbf{L}_2$, $\mathbf{L}_3$, and $\mathbf{L}_4$ can be obtained by inspection from Eq. (2).

A lamina under consideration is transversely isotropic, so the stresses are related to strains by

$$\sigma=\overline{\mathbf{Q}}\boldsymbol{\varepsilon}, \tag{3}$$

where $\boldsymbol{\sigma}=[\sigma_z\sigma_\theta\sigma_r\tau_{r\theta}\tau_{rz}\tau_{z\theta}]^T$ is the vector of stresses and

$$\overline{\mathbf{Q}}=\begin{bmatrix} \overline{Q}_{11} & \overline{Q}_{12} & \cdots & \overline{Q}_{16} \\ & \overline{Q}_{22} & \cdots & \overline{Q}_{26} \\ & & \ddots & \vdots \\ \text{sym.} & & & \overline{Q}_{66} \end{bmatrix} \tag{4}$$

is the matrix of the off-principal-axis stiffness coefficients of the lamina whose expressions in terms of engineering constants are given by Vinson and Sierakowski (1987).

With the above basic equations, we apply the Hamilton's principle to derive the dispersion equation for the shell. In this study, we first use finite elements to model the radial displacement of the shell and then employ the complex exponentials to model the axial and circumferential displacements. In view of the heterogeneity of the laminated composite shell in the radial direction, an annular element shown in Fig. 2 is used in the subdivision of the shell so as to achieve high computational precision. The element has the inner, middle, and outer nodal surfaces, *i*, *m*, *o*, and each nodal surface has three degrees of freedom, *u*, *v*, *w*. Hence the vector of the unknown displacement amplitudes of the element is expressed as $\mathbf{V}^e=[u_iv_iw_iu_mv_mw_mu_ov_ow_o]^T$. Suppose that the shell is subdivided into N elements in the radial direction, and that r_{j-1} and r_j represent, respectively, the inner and outer radii of any element *j*. The displacements within an element are thus approximated as

$$\mathbf{u}=\mathbf{N}(r)\mathbf{V}^e, \tag{5}$$

where

$$\mathbf{N}(r)=[(1-3\hat{r}+2\hat{r}^2)\mathbf{I}4(\hat{r}-\hat{r}^2)\mathbf{I}(-\hat{r}+2\hat{r}^2)\mathbf{I}] \tag{6}$$

is the shape function matrix of the element. Here $\hat{r}=(r-r_{j-1})/(r_j-r_{j-1})$, $r_{j-1}\leqslant r\leqslant r_j$ and $\mathbf{I}$ is a 3×3 identity matrix.

The potential energy of the element in the absence of body forces is given by

$$V=\int_{-\infty}^{+\infty}\int_0^{2\pi}\int_{r_{j-1}}^{r_j}\boldsymbol{\varepsilon}^T\boldsymbol{\sigma}r\,dr\,d\theta\,dz. \tag{7}$$

Invoking Eqs. (1) through (3), we obtain

$$\begin{aligned} V=\int_{-\infty}^{+\infty}\int_0^{2\pi}\int_{r_{j-1}}^{r_j}\Bigg(&\frac{\partial\mathbf{u}^T}{\partial z}\mathbf{D}_{11}\frac{\partial\mathbf{u}}{\partial z}+\frac{1}{r}\frac{\partial\mathbf{u}^T}{\partial z}\mathbf{D}_{12}\frac{\partial\mathbf{u}}{\partial\theta} \\ &+\frac{\partial\mathbf{u}^T}{\partial z}\mathbf{D}_{13}\frac{\partial\mathbf{u}}{\partial r}+\frac{1}{r}\frac{\partial\mathbf{u}^T}{\partial z}\mathbf{D}_{14}\mathbf{u}+\frac{1}{r}\frac{\partial\mathbf{u}^T}{\partial\theta}\mathbf{D}_{12}^T\frac{\partial\mathbf{u}}{\partial z} \\ &+\frac{1}{r^2}\frac{\partial\mathbf{u}^T}{\partial\theta}\mathbf{D}_{22}\frac{\partial\mathbf{u}}{\partial\theta}+\frac{1}{r}\frac{\partial\mathbf{u}^T}{\partial\theta}\mathbf{D}_{23}\frac{\partial\mathbf{u}}{\partial r}+\frac{1}{r^2}\frac{\partial\mathbf{u}^T}{\partial\theta}\mathbf{D}_{24}\mathbf{u} \\ &+\frac{\partial\mathbf{u}^T}{\partial r}\mathbf{D}_{13}^T\frac{\partial\mathbf{u}}{\partial z}+\frac{1}{r}\frac{\partial\mathbf{u}^T}{\partial r}\mathbf{D}_{23}^T\frac{\partial\mathbf{u}}{\partial\theta}+\frac{\partial\mathbf{u}^T}{\partial r}\mathbf{D}_{33}\frac{\partial\mathbf{u}}{\partial r} \\ &+\frac{1}{r}\frac{\partial\mathbf{u}^T}{\partial r}\mathbf{D}_{34}\mathbf{u}+\frac{1}{r}\mathbf{u}^T\mathbf{D}_{14}^T\frac{\partial\mathbf{u}}{\partial z}+\frac{1}{r^2}\mathbf{u}^T\mathbf{D}_{24}^T\frac{\partial\mathbf{u}}{\partial\theta} \\ &+\frac{1}{r}\mathbf{u}^T\mathbf{D}_{34}^T\frac{\partial\mathbf{u}}{\partial r}+\frac{1}{r^2}\mathbf{u}^T\mathbf{D}_{44}\mathbf{u}\Bigg)r\,dr\,d\theta\,dz, \end{aligned} \tag{8}$$

where $\mathbf{D}_{ij}=\mathbf{L}_i^T\overline{\mathbf{Q}}\mathbf{L}_j$ $(i,j=1,2,3,4)$.

The kinetic energy of the element is expressed in terms of the displacement vector as

$$T=\int_{-\infty}^{+\infty}\int_0^{2\pi}\int_{r_{j-1}}^{r_j}\frac{\partial\mathbf{u}^T}{\partial t}\frac{\partial\mathbf{u}}{\partial t}\rho r\,dr\,d\theta\,dz, \tag{9}$$

where ρ is the mass density of the material of the element.

Applying the Hamilton's principle gives the characteristic equations of the element as follows:

$$-\mathbf{A}_1^e\frac{\partial^2\mathbf{V}^e}{\partial z^2}-\mathbf{A}_2^e\frac{\partial^2\mathbf{V}^e}{\partial\theta\partial z}-\mathbf{A}_3^e\frac{\partial^2\mathbf{V}^e}{\partial\theta^2}+\mathbf{A}_4^e\frac{\partial\mathbf{V}^e}{\partial z}+\mathbf{A}_5^e\frac{\partial\mathbf{V}^e}{\partial\theta}+\mathbf{A}_6^e\mathbf{V}^e+\mathbf{M}^e\frac{\partial^2\mathbf{V}^e}{\partial t^2}=0, \tag{10}$$

where

$$\mathbf{A}_1^e=\int_{r_{j-1}}^{r_j}\mathbf{N}^T\mathbf{D}_{11}\mathbf{N}r\,dr, \tag{11}$$

$$\mathbf{A}_2^e=\int_{r_{j-1}}^{r_j}\mathbf{N}^T(\mathbf{D}_{12}+\mathbf{D}_{12}^T)\mathbf{N}\,dr, \tag{12}$$

$$\mathbf{A}_3^e=\int_{r_{j-1}}^{r_j}\frac{1}{r}\mathbf{N}^T\mathbf{D}_{22}\mathbf{N}\,dr, \tag{13}$$

$$\mathbf{A}_4^e=\int_{r_{j-1}}^{r_j}\left[-\mathbf{N}^T\mathbf{D}_{13}\frac{d\mathbf{N}^T}{dr}+\frac{d\mathbf{N}^T}{dr}\mathbf{D}_{13}^T\mathbf{N}+\frac{1}{r}\mathbf{N}^T(\mathbf{D}_{14}^T-\mathbf{D}_{14})\mathbf{N}\right]r\,dr, \tag{14}$$

$$\mathbf{A}_5^e=\int_{r_{j-1}}^{r_j}\left[-\mathbf{N}^T\mathbf{D}_{23}\frac{d\mathbf{N}^T}{dr}+\frac{d\mathbf{N}^T}{dr}\mathbf{D}_{23}^T\mathbf{N}+\frac{1}{r}\mathbf{N}^T(\mathbf{D}_{24}^T-\mathbf{D}_{24})\mathbf{N}\right]dr, \tag{15}$$

$$\mathbf{A}_6^e=\int_{r_{j-1}}^{r_j}\left(\frac{d\mathbf{N}^T}{dr}\mathbf{D}_{33}\frac{d\mathbf{N}}{dr}+\frac{1}{r}\frac{d\mathbf{N}^T}{dr}\mathbf{D}_{34}\mathbf{N}+\frac{1}{r}\mathbf{N}^T\mathbf{D}_{34}^T\frac{d\mathbf{N}}{dr}+\frac{1}{r^2}\mathbf{N}^T\mathbf{D}_{44}\mathbf{N}\right)r\,dr, \tag{16}$$

$$\mathbf{M}^e=\int_{r_{j-1}}^{r_j}\mathbf{N}^T\mathbf{N}\rho r\,dr. \tag{17}$$

The axial and circumferential displacement vector $\mathbf{V}^e(\theta,z,t)$ can be expanded as the complex exponentials

$$\mathbf{V}^e(\theta,z,t)=\mathbf{U}^e\exp i(n\theta+k_z z-\omega t), \tag{18}$$

where ω is the circular frequency, n is the wave number in the circumferential direction, and $k_z=k\cos\beta$ is the wave number in the z-axis direction. When a helical wave of wave number k propagates in the shell at an arbitrary β angle with respect to the z-axis, we have

$$n=R_o k\sin\beta,\quad k_z=k\cos\beta. \tag{19}$$

Substituting Eq. (18) into Eq. (10), we find

$$[\mathbf{K}^e-\omega^2\mathbf{M}^e]\mathbf{U}^e=\mathbf{0}, \tag{20}$$

where

$$\mathbf{K}^e=\mathbf{A}_1^e k_z^2+\mathbf{A}_2^e n k_z+\mathbf{A}_3^e n^2+i\mathbf{A}_4^e k_z+i\mathbf{A}_5^e n+\mathbf{A}_6^e. \tag{21}$$

Assembling elements at the nodal surfaces, we find the dispersion equation for the shell

$$[\mathbf{K}-\omega^2\mathbf{M}]\mathbf{U}=\mathbf{0}. \tag{22}$$

During the process of assembling elements, we make use of the following interface and boundary conditions:

$$\mathbf{p}_1^i=0, \tag{23}$$

$$\mathbf{p}_j^o=\mathbf{p}_{j+1}^i\quad\mathbf{U}_j^o=\mathbf{U}_{j+1}^i\quad\text{for } 1<j<N-1, \tag{24}$$

$$\mathbf{p}_N^o=0, \tag{25}$$

in which $\mathbf{p}$ is the vector of the surface tractions, the subscripts denote the element numbers, and the superscripts denote the inner and outer surfaces of the element.

When the wave number is specified, we can solve Eq. (22) for the circular frequency of the shell, and accordingly obtain the relationship between the wave number and circular frequency, that is, the dispersion relationship for the shell.

III. WAVE SURFACES

Unlike isotropic cases, wave propagation in anisotropic media is complicated by both dispersion and anisotropic nature of waves. To visualize the effects of anisotropy on wave propagation, Liu *et al.* (1991) introduced a set of six characteristic wave surfaces: phase velocity surface (PVS), phase slowness surface (PSS), phase wave surface (PWS), group velocity surface (GVS), group slowness surface (GSS), and group wave surface (GWS). The PVS shows the dependence of the phase velocity of a plane wave on the direction of wave propagation. The PSS shows the dependence of the relative arrival time of a plane wave on the direction of wave propagation; a normal vector on this surface is proportional to a position vector of the wave front from a point wave source. The PWS shows the phase wave front emanating from a point wave source. The GVS shows the dependence of the energy propagation velocity of a plane wave on the direction of wave propagation. The GSS shows the dependence of the relative arrival time of the energy of a plane wave on the direction of wave propagation; a normal vector on this surface is proportional to a vector that represents the direction of energy flow from a point wave source. The GWS shows the group wave front emanating from a point wave source. In what follows, we will derive the mathematical expressions for these characteristic wave surfaces with the aid of the Rayleigh quotient concept.

The circular frequency for the mth mode can be written in terms of the Rayleigh quotient as

$$\omega_m^2=\frac{\varphi_m^L\mathbf{K}\varphi_m^R}{\varphi_m^L\mathbf{M}\varphi_m^R}, \tag{26}$$

where φ_m^L and φ_m^R are the mth transposed left and right eigenvectors of Eq. (22).

The phase velocity of the mth mode of the plane wave is defined as

$$c_m=\omega_m/k \tag{27}$$

or

$$\omega_m=c_m k. \tag{28}$$

Considering Eq. (26), we obtain the phase velocity surface for the mth mode,

$$f_c(c_m,\beta)=\varphi_m^L\mathbf{K}\varphi_m^R-c_m^2k^2\varphi_m^L\mathbf{M}\varphi_m^R=0. \tag{29}$$

The phase slowness for the mth mode is defined as the reciprocal of the phase velocity

$$s_m = 1/c_m . \tag{30}$$

From Eqs. (26) and (27), we have the phase slowness surface for the mth mode,

$$f_s(s_m, \beta) = s_m^2 \varphi_m^L \mathbf{K} \varphi_m^R - k^2 \varphi_m^L \mathbf{M} \varphi_m^R = 0. \tag{31}$$

The phase wave surface for the mth mode is defined as

$$\mathbf{r}_m = c_m \nabla \mathbf{f}_s / (\mathbf{e}_s \cdot \nabla \mathbf{f}_s), \tag{32}$$

where

$$\nabla \mathbf{f}_s = f_{s,s_m} \mathbf{e}_s + f_{s,\beta} / s_m \mathbf{e}_n . \tag{33}$$

$\mathbf{e}_s = [\cos\beta \sin\beta]$ and $\mathbf{e}_n$ are, respectively, the normal and tangential unit vectors to the surface traced by the tip of the vector $\mathbf{r}_m$. Substituting Eq. (33) into Eq. (32), we find

$$\mathbf{r}_m = c_m \mathbf{e}_s + c_m^2 f_{s,\beta} / f_{s,s_m} \mathbf{e}_n . \tag{34}$$

Hence

$$|\mathbf{r}_m| = c_m \sqrt{1 + \alpha_s^2}, \quad \psi_s = \tan^{-1} \alpha_s , \tag{35}$$

where

$$\alpha_s = c_m f_{s,\beta} / f_{s,s_m}. \tag{36}$$

Differentiating Eq. (31) with respect to s_m and β, respectively, we obtain

$$f_{s,s_m} = 2 \varphi_m^L \mathbf{K} \varphi_m^R / c_m , \tag{37}$$

$$f_{s,\beta} = \varphi_m^L \mathbf{K}_{,\beta} \varphi_m^R / c_m^2, \tag{38}$$

where

$$\mathbf{K}_{,\beta} = -\mathbf{A}_1 k^2 \sin 2\beta + \mathbf{A}_2 iRk^2 \cos 2\beta + \mathbf{A}_3 R^2 k^2 \sin 2\beta - \mathbf{A}_4 ik \sin\beta + \mathbf{A}_5 Rk \cos\beta. \tag{39}$$

The phase wave surface represents the locus of the tips of the vector $\mathbf{r}_m$ and it makes an angle $\phi = \psi_s + \beta$ with the z-axis.

The group velocity represents the rate at which energy is transported. It is defined as

$$c_g = d\omega / dk. \tag{40}$$

Differentiating Eq. (26) with respect to k, we obtain the group velocity for the mth mode,

$$c_{gm} = \frac{\varphi_m^L \mathbf{K}_{,k} \varphi_m^R}{2 \omega_m \varphi_m^L \mathbf{M} \varphi_m^R}, \tag{41}$$

where

$$\mathbf{K}_{,k} = 2\mathbf{A}_1 k \cos^2\beta + \mathbf{A}_2 iR \sin 2\beta + 2\mathbf{A}_3 R^2 k \sin^2\beta + i\mathbf{A}_4 \cos\beta + \mathbf{A}_5 R \sin\beta. \tag{42}$$

From Eq. (41), we have the group velocity surface for the mth mode,

$$f_g(c_{gm}, \beta) = \varphi_m^L \mathbf{K}_{,k} \varphi_m^R - 2 c_{gm} \omega_m \varphi_m^L \mathbf{M} \varphi_m^R . \tag{43}$$

The group slowness for the mth mode is defined as the reciprocal of the group velocity

$$q_m = 1/c_{gm} . \tag{44}$$

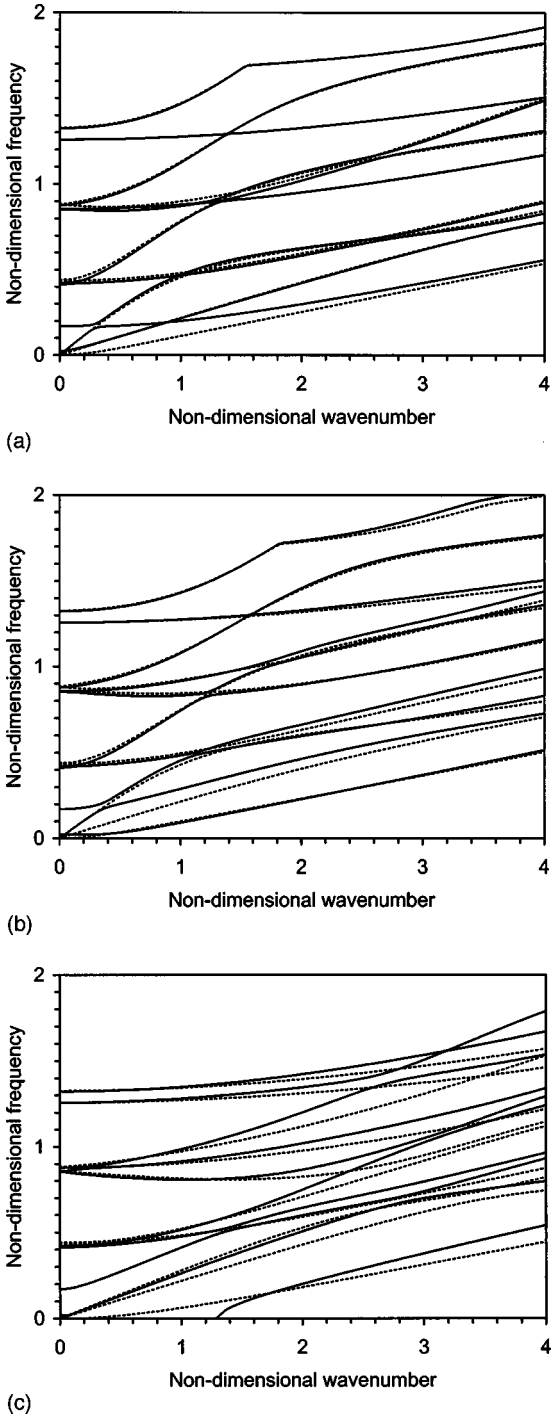


FIG. 3. Dispersion curves for waves propagating in $(C0/G \pm 45)_s$ circular cylindrical shells. (a) $\beta=0$; (b) $\beta=30°$; (c) $\beta=90°$. —— $R=1$; - - - $R=100$.

From Eq. (43), we obtain the group slowness surface for the mth mode,

$$f_q(q_m, \beta) = q_m \varphi_m^L \mathbf{K}_{,k} \varphi_m^R - 2 \omega_m \varphi_m^L \mathbf{M} \varphi_m^R . \tag{45}$$

The group wave surface for the mth mode is defined as

$$\mathbf{R}_m = c_m \nabla \mathbf{f}_q / (\mathbf{e}_q \cdot \nabla \mathbf{f}_q), \tag{46}$$

where

$$\nabla \mathbf{f}_q = f_{q,q_m} \mathbf{e}_g + f_{q,\beta} / q_m \mathbf{e}_n . \tag{47}$$

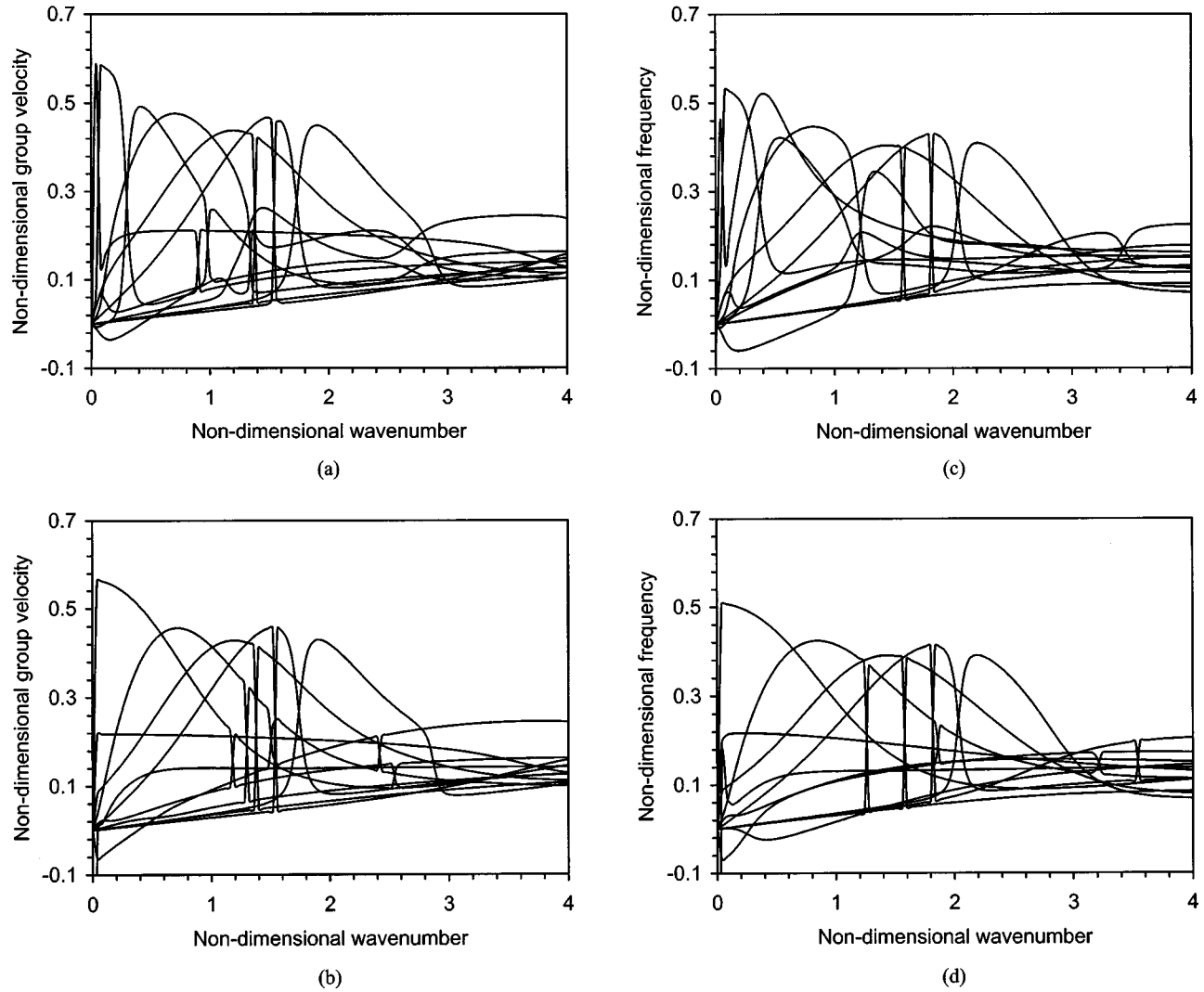


FIG. 4. Group velocity spectra for waves propagating in $(C0/G\pm 45)_s$ circular cylindrical shells. (a) $\beta=0$, $\bar{R}=1$; (b) $\beta=0$, $\bar{R}=100$; (c) $\beta=30°$, $\bar{R}=1$; (d) $\beta=30°$, $\bar{R}=100$.

$\mathbf{e}_q=[\cos\beta \sin\beta]$ and $\mathbf{e}_n$ are, respectively, the normal and tangential unit vectors to the group wave surface. Substituting Eq. (47) into Eq. (46), we find

$$\mathbf{R}_m=c_{gm}\mathbf{e}_q+c_{gm}^2 f_{q,\beta}/f_{q,q_m}\mathbf{e}_n\,. \tag{48}$$

Hence

$$|\mathbf{R}_m|=c_{gm}\sqrt{1+\alpha_q^2},\quad \psi_q=\tan^{-1}\alpha_q\,, \tag{49}$$

where

$$\alpha_q=c_{gm}f_{q,\beta}/f_{q,q_m}. \tag{50}$$

Differentiating Eq. (45) with respect to q_m and β, respectively, we have

$$f_{q,q_m}=\varphi_m^L\mathbf{K},_k\varphi_m^R, \tag{51}$$

$$f_{q,\beta}=q_m\varphi_m^L\mathbf{K},_{k\beta}\varphi_m^R-\varphi_m^L\mathbf{K},_\beta\varphi_m^R/\omega_m\,, \tag{52}$$

where

$$\mathbf{K},_{k\beta}=-2\mathbf{A}_1 k\sin 2\beta+2\mathbf{A}_2 iR\cos 2\beta+2\mathbf{A}_3R^2k\sin 2\beta - i\mathbf{A}_4\cos\beta-i\mathbf{A}_5R\sin\beta. \tag{53}$$

The group wave surface represents the locus of the endpoints of the vector $\mathbf{R}_m$ and the angle of the vector to the z axis is $\phi=\psi_q+\beta$.

Until now, six representation wave surfaces have been obtained, of which the most important surface is the GWS, because it represents the energy surface of propagation from a point source wave. The slowness diagrams are useful to illustrate reflections and refractions of waves.

IV. NUMERICAL RESULTS AND DISCUSSION

In this section, numerical results are presented for dispersion and characteristic surfaces of waves in laminated composite cylindrical shell. In laminate codes used below, a lamina numbering increases from the inner to outer surface; letters C and G represent carbon/epoxy and glass/epoxy, respectively; the number following the letters indicates the azimuthal angle of the fiber orientation with respect to the z axis; the subscript s denotes that the multilayered shell is symmetrically stacked about the middle surface. The mate-

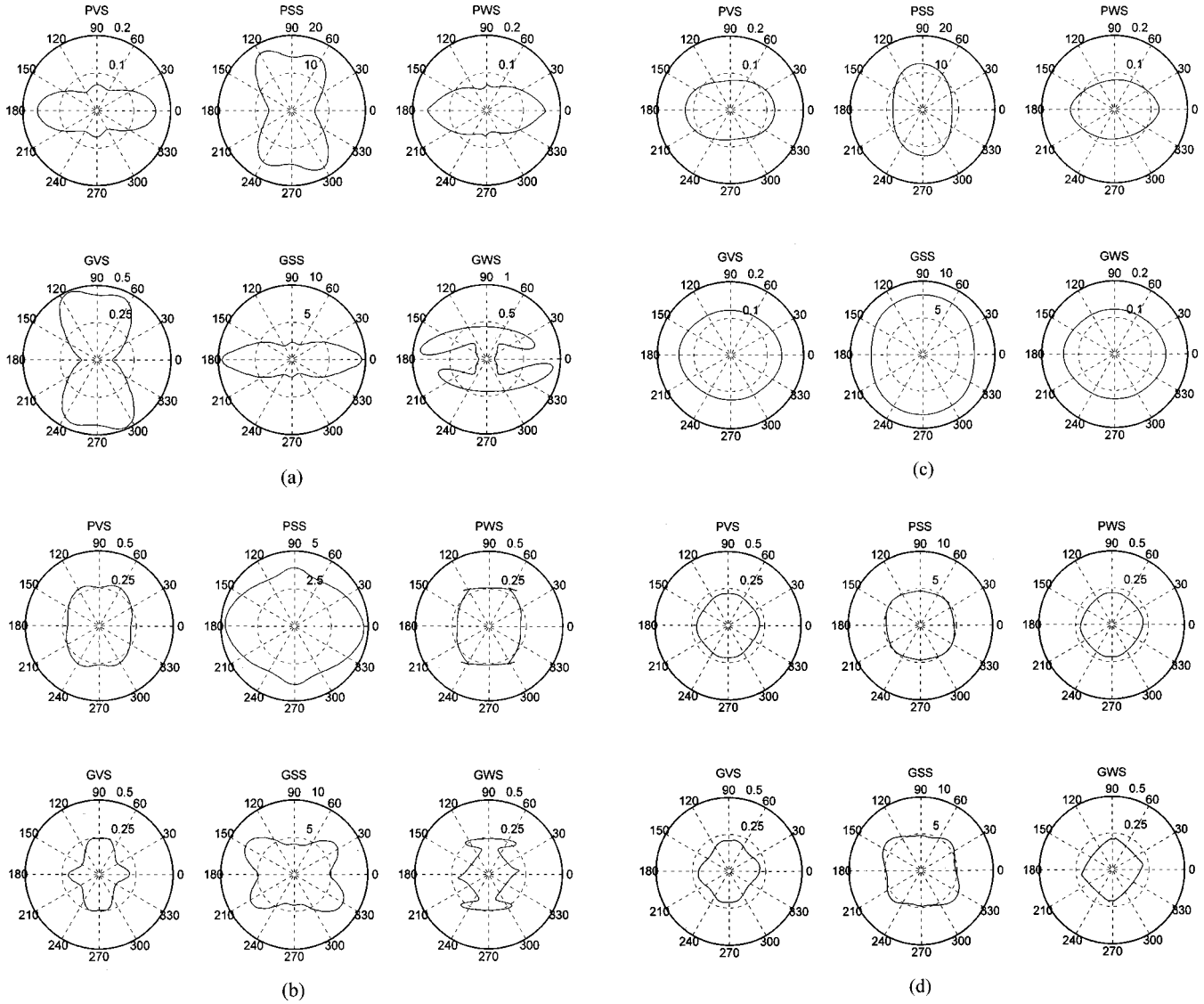


FIG. 5. Characteristic wave surfaces for $(C0/G\pm45)_s$ circular cylindrical shells. (a) The first mode for $\bar{R}=1$; (b) the second mode for $\bar{R}=1$; (c) the first mode for $\bar{R}=100$; (d) the second mode for $\bar{R}=100$.

rial properties of the shell are taken from Takahashi and Chou (1987). The following dimensionless parameters are adopted:

$$\bar{k}=k(R_o-R_i),\quad \bar{\lambda}=\lambda/(R_o-R_i),\quad \bar{R}=R_i/(R_o-R_i),$$
$$\bar{\omega}=\omega(R_o-R_i)\sqrt{\rho/Q_{11}},$$

where the reference properties Q_{11} and ρ are the Young's modulus in the fiber direction and mass density of $C0$. Since the present method is based on the theory of three-dimensional elasticity, it is applicable to both a circular cylindrical shell and a circular cylinder. The distinction between waves in a cylindrical shell and waves in a cylinder can be explored through studying the effect of the ratio of the inner radius to the thickness of the shell. To this end, two ratios of radius to thickness $\bar{R}=1$ and 100 are used here. The former is for a cylinder, while the latter is for a cylindrical shell.

We first discuss the dispersion behaviors of waves. In Figs. 3(a)–(c), the dispersion curves for a $(C0/G\pm45)_s$ shell are compared with those for a $(C0/G\pm45)_s$ cylinder. The propagation directions of waves are chosen as $\beta=0$, 30 and 90°, respectively. They are respectively referred to as the axial, helical, and circumferential waves. The distinction between the axial waves in the shell and cylinder can be found from Fig. 3(a). The only curve for the first propagation mode in the case of the shell is straight lines passing through the origin. This reveals that the wave propagation for the first mode for the shell is nondispersive and all frequencies have the same phase velocity, whereas the wave propagation for higher order modes is dispersive. The wave propagation for all the modes in the case of the cylinder is dispersive. In addition, the ratio of radius to thickness affects only those curves for lower order modes and smaller wave numbers. From Fig. 3(b) it can be seen that for waves propagating at an angle $\beta=30°$ with respect to the axial direction, the ratio of radius to thickness also affects only the curves for lower order modes, but this effect increases as the wave number increases. The results in Fig. 3(c) indicate that the ratio of

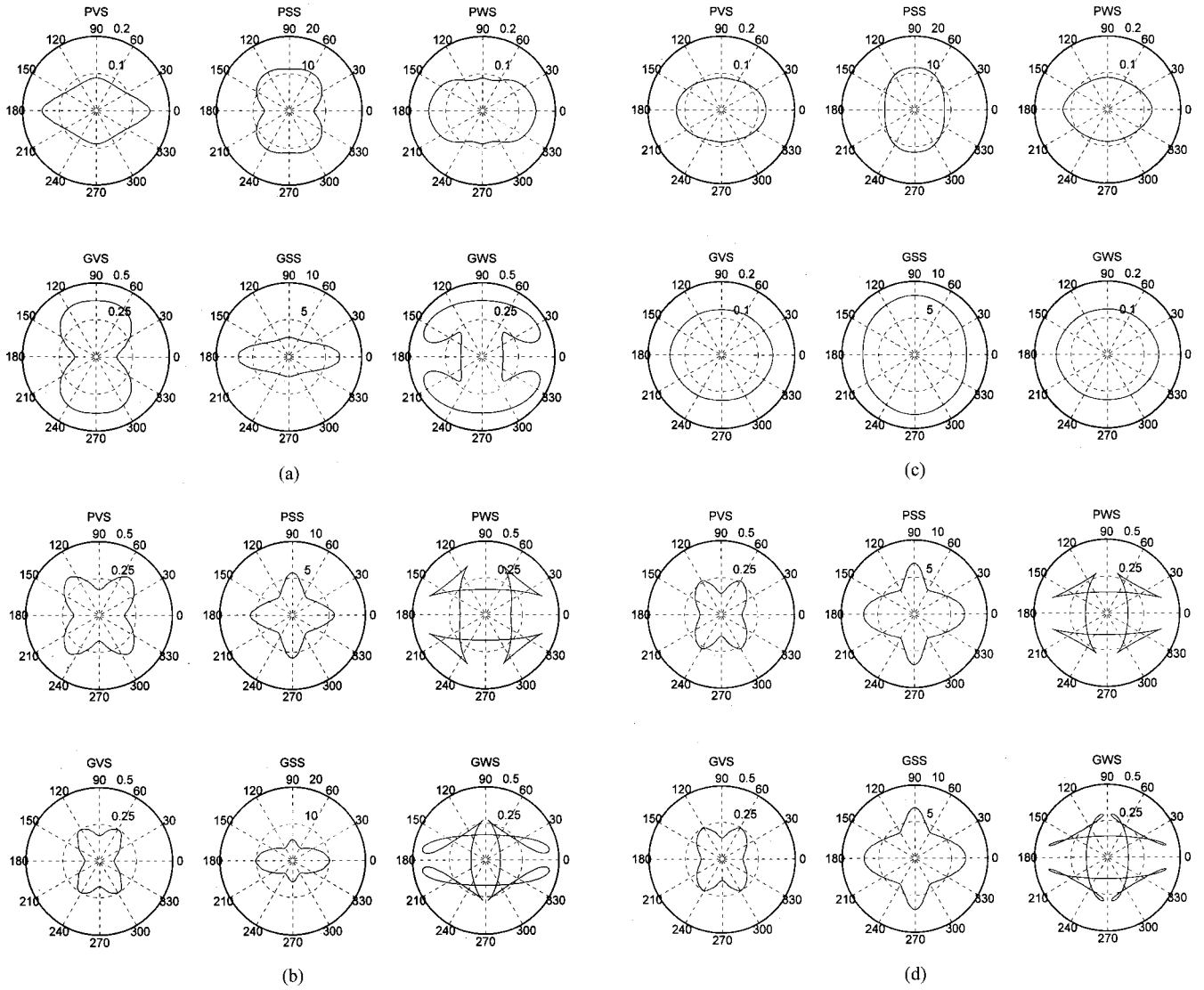


FIG. 6. Same as Fig. 6 but for $(C0/G90/G0)_s$ circular cylindrical shells.

radius to thickness has considerable influence on the curves for all the modes, and that the larger the wave number, the stronger this effect.

We now turn to the computation of group velocity spectra. In Figs. 4(a)–(d), the group velocity spectra for a $(C0/G\pm45)_s$ cylindrical shell are compared with those for a $(C0/G\pm45)_s$ cylinder. The propagation directions of waves are chosen as $\beta=0$ and 30°, respectively. Similar to phenomena observed by Liu *et al.* (1991) for laminated composite plates, negative group velocities occur at a range of smaller wave numbers for both the shell and cylinder. The range of wave numbers producing the negative group velocities for the wave with $\beta=30°$ is larger than that for the axial wave. This indicates that the range varies with the wave normal. The other difference between the shell and cylinder waves lies in the fact that the curves for the shell present more abrupt changes than those for the cylinder. It is thus evident that the ratio of radius to thickness can alter the pattern of the group velocity spectra. Besides, it is very clear that the velocities of the energy flux depend considerably on the wave number and propagation modes of waves.

Now, let us look at characteristic wave surfaces. Figures 5(a)–(d) illustrate six representation surfaces for the first two propagation modes for a $(C0/G\pm45)_s$ cylindrical shell and a $(C0/G\pm45)_s$ cylinder. From these figures, we can see very clearly that the wave propagation in the shells depends strongly upon not only the wave propagation modes but also the anisotropy of composites and the ratio of radius to thickness. Energy flux deviation is observed for all waves as none of the slowness curves are circular. Figures 6(a)–(d) are the same as Figs. 5(a)–(d) but for a $(C0/G90/G0)_s$ shell and a $(C0/G90/G0)_s$ cylinder. These results can further confirm the above observations. On the other hand, comparing Figs. 5(a)–(d) with Figs. 6(a)–(d) reveals that characteristic wave surfaces vary with the lay-ups of the laminated composite shells.

V. CONCLUSIONS

The frequency spectra and group velocity spectra and characteristic surfaces of waves in laminated composite circular cylindrical shell have been investigated. The method of

approach is formulated within the framework of the theory of three-dimensional elasticity, and is thus accurate in comparison with ones using various shell theories. The use of finite elements through the radial direction makes it easier to deal with composite circular cylindrical shells made of an arbitrary number of anisotropic layers, of arbitrary lay-ups, and of any type of materials. Furthermore, the method is capable of reducing the spatial dimensions of a problem by one and of omitting tedious preprocessors occupying a substantial part of finite element methods, and accordingly of reducing a great deal of computational labor. Based on the calculated results, the following conclusions can be drawn:

(a) The ratio of radius to thickness has stronger influence on the frequency spectra in the circumferential wave than on that in the axial wave.
(b) The ratio of radius to thickness can alter the pattern of the group velocity spectra.
(c) The negative group velocity appears at a range of smaller wave numbers and the range varies as the wave normal.
(d) The characteristic wave surfaces vary with the propagation modes of waves, the ratio of radius to thickness, and the lay-ups of laminated composite cylindrical shells.

Datta, S. K., Ledbetter, H. M., Shindo, Y., and Shah, A. H. (**1988a**). ''Phase velocity and attenuation of plane elastic waves in a particle-reinforced composite medium,'' Wave Motion **10**, 171–182.

Datta, S. K., Shah, A. H., Bratton, R. L., and Chakraborty, T. (**1988b**). ''Wave propagation in laminated composite plates,'' J. Acoust. Soc. Am. **83**, 2020–2026.

Han, X., Liu, G. R., Xi, Z. C., and Lam, K. Y. (**1999**). ''Transient waves in a functionally graded cylinder,'' Int. J. Solids Struct. (to be published).

Huang, K. H., and Dong, S. B. (**1984**). ''Propagating waves and edge vibrations in anisotropic composite cylinders,'' J. Sound Vib. **96**, 363–379.

Kline, R. A. (**1992**). *Nondestructive Characterization of Composite Media* (Technomic, Lancaster).

Liu, G. R., Tani, J., Ohyoshi, T., and Watanabe, K. (**1991**). ''Characteristic wave surfaces in anisotropic laminated plates,'' J. Vibr. Acoust. **113**, 279–285.

Mal, A. K. (**1988**). ''Wave propagation in layered composite laminates under periodic surface loads,'' Wave Motion **10**, 257–266.

Markus, S., and Mead, D. J. (**1995a**). ''Axisymmetric and asymmetric wave motion in orthotropic cylinders,'' J. Sound Vib. **181**, 127–147.

Markus, S., and Mead, D. J. (**1995b**). ''Wave motion in a 3-layered, orthotropic isotropic orthotropic composite shell,'' J. Sound Vib. **181**, 149–167.

Nayfeh, A. H. (**1995**). *Wave Propagation in Layered Anisotropic Media with Applications to Composites* (Elsevier, Amsterdam).

Nelson, R. B., Dong, S. B., and Kalra, R. D. (**1971**). ''Vibrations and waves in laminated orthotropic circular cylinders,'' J. Sound Vib. **18**, 429–444.

Rattanwangcharoen, N., Shah, A., and Datta, S. K. (**1994**). ''Reflection of waves at the free edge of a laminated circular cylinder,'' ASME J. Appl. Mech. **61**, 323–329.

Rattanwangcharoen, N., Zhuang, W., Shah, A., and Datta, S. K. (**1997**). ''Axisymmetric guided waves in jointed laminated cylinders,'' ASCE J. Eng. Mech. **123**, 1020–1026.

Rose, J. L. (**1999**). *Ultrasonic Waves in Solid Media* (Cambridge University Press, Cambridge, UK).

Takahashi, K., and Chou, Tsu-Wei (**1987**). ''Non-linear deformation and failure behavior of Carbon/Glass hybrid laminates,'' J. Compos. Mater. **21**, 396–407.

Vinson, J. R., and Sierakowski, R. L. (**1987**). *The Behavior of Structures Composed of Composite Materials* (Martinus Nijhoff, Dordrecht).

Xi, Z. C., Liu, G. R., Lam, K. Y., and Shang, H. M. (**2000a**). ''Strip element method for analyzing wave scattering by a crack in a fluid-filled laminated composite shell,'' Compos. Sci. Technol. **60**, 1985–1996.

Xi, Z. C., Liu, G. R., Lam, K. Y., and Shang, H. M. (**2000b**). ''Strip element method for analyzing wave scattering by a crack in an immersed laminated composite cylinder,'' J. Acoust. Soc. Am. **108**, 175–183.

Xi, Z. C., Liu, G. R., Lam, K. Y., and Shang, H. M. (**2000c**). ''Strip element method for analyzing wave scattering by a crack in a laminated composite cylinder,'' ASME J. Appl. Mech. **67**, 427–429.

Xi, Z. C., Yam, L. H., and Leung, T. P. (**1996**). ''Semi-analytical study of free vibration of composite shells of revolution based on the Reissner-Mindlin assumption,'' Int. J. Solids Struct. **33**, 851–863.

Xi, Z. C., Yam, L. H., and Leung, T. P. (**1997a**). ''Free vibration of a laminated composite circular cylindrical shell partially filled with a fluid,'' Composites, Part B **28B**, 359–375.

Xi, Z. C., Yam, L. H., and Leung, T. P. (**1997b**). ''Free vibration of a partially fluid-filled cross-ply laminated composite circular cylindrical shell,'' J. Acoust. Soc. Am. **101**, 909–917.

Xi, Z. C., Yam, L. H., and Leung, T. P. (**1999**). ''Nonlinearly elastic free vibration of a laminated composite shell of revolution,'' Int. J. Mech. Sci. **41**, 649–661.

Yuan, F. G., and Hsieh, C. C. (**1998**). ''Three-dimensional wave propagation in composite cylindrical shell,'' Compos. Struct. **42**, 153–167.

Zhuang, W., Shah, A., and Datta, S. K. (**1997**). ''Axisymmetric guided wave scattering by cracks in welded steel pipes,'' J. Pressure Vessel Technol. **119**, 401–406.

Analysis of the acoustic scattering at variable incidences from an extra thin cylindrical shell bounded by hemispherical endcaps

N. Touraine, L. Haumesser, D. Décultot, and G. Maze[a)]
Laboratoire d'Acoustique Ultrasonore et d'Electronique, LAUE, UPRESA CNRS 6068, Université du Havre, Place Robert Schuman, 76610 Le Havre, France

A. Klauson and J. Metsaveer
Tallinn Technical University, Department of Mechanics, Ehitajate tee 5, 19086 Tallinn, Estonia

(Received 27 January 2000; revised 13 June 2000; accepted 16 June 2000)

Through an experimental approach, in this paper we investigate the acoustic wave scattering processes involved in the acoustic backscattering at variable incidences from an air-filled submerged cylindrical shell with hemispherical endcaps. Given the 1% shell thickness and the explored low frequency domain, the wave types studied are the circumferential or helical S_0 wave and the helical T_0 wave only. Between the axial (in the direction of the main axis of the object) and the normal incidences (normal to the main axis), two distinct angular zones can be observed depending on hemispherical or cylindrical excitation. In these zones, after a pressure wave excitation, different series of echoes on the echo wave forms are identified by their arrival times and related wave types. From results in the time domain and those obtained in the frequency domain, each acoustic response from the target corresponding to the two zones of excitation is compared with the acoustic response of canonical objects (spherical shell for axial excitation and tube for normal excitation). This analysis of the acoustic response from the target at various incidences, highlights the influence of both the endcaps and the finite length for a cylindrical shell on scattering. The study is intended to make a contribution to the knowledge of the identification of such geometrically complex objects.
 [S0001-4966(00)03909-6]

PACS numbers: 43.40.Fz, 43.40.Rj, 43.30.Gv [CBB]

I. INTRODUCTION

The increasing interest for the detection of targets having complex geometry has attracted the attention of many authors during the last decade. These authors often considered, in a preliminary approach, only particular directions of observation. Among studies by these researchers, numerical and/or experimental cases of axial[1–5] and normal[1,6] incidences on finite cylindrical shells bounded by hemispherical endcaps are of particular interest. However, in the axial incidence (excitation on one of the hemispherical caps), numerical results are often difficult to verify experimentally. Indeed for a given shell thickness and in an appropriate frequency window, the *A* wave (Scholte–Stoneley) is generated,[7,8] but its experimental observation on this type of target is difficult to confirm. In fact the *A* wave is particularly sensitive to variations of shell thickness due to internal solders introduced during the assembling of hemispherical and cylindrical parts.[9] Furthermore the sound scattering by this complex object is insufficiently explored for variable incidences between axial and normal positions. Certain studies provide in some detail results in oblique incidence for infinite[10,11] cylindrical shells and finite[10,12] ones closed on both ends by flat disks. At such incidences the resonances due to surface helical waves have been observed and identified.[13–15] In a low frequency range, Rumerman[16] studies theoretically the acoustic scattering from a finite cylindrical shell and interprets his results in terms of resonances associated with the shell length interacting with the circumferential resonances of the infinite shell. In this study the author does not take into account the effects of endcaps shape in the interpretation of results by the proposed mathematical model. Bao and Überall[17] investigate, both in time and frequency domains, the acoustic scattering from a cylinder bounded by hemispheres at variable incidences. The authors describe wave traveling paths on the surface of such an object. In order to experimentally validate his theoretical approach at oblique incidence, Bao[18] applies the model on a simpler target, a finite cylindrical shell closed with disks, to estimate the time response of echoes reradiated by the object.

In light of these results, our aim in this paper is to examine in detail the response from a cylindrical shell bounded by hemispherical shell caps after a short pulse excitation at any incidence, and to interpret the mechanism of wave propagation in the time domain, and the resonant phenomena in the considered frequency range. In order to achieve this study and to avoid the influence of the *A* wave, an extra thin shell is used.

In Sec. II the authors outline geometric characteristics of the studied object, provide experimental conditions and present an overview of results. The case where the target is excited on one of the hemispherical caps is treated in Sec. III. Numerical results for the axisymmetric case are used to provide the identification of the nature of the propagating

a)Electronic mail: maze@iut.univ-lehavre.fr

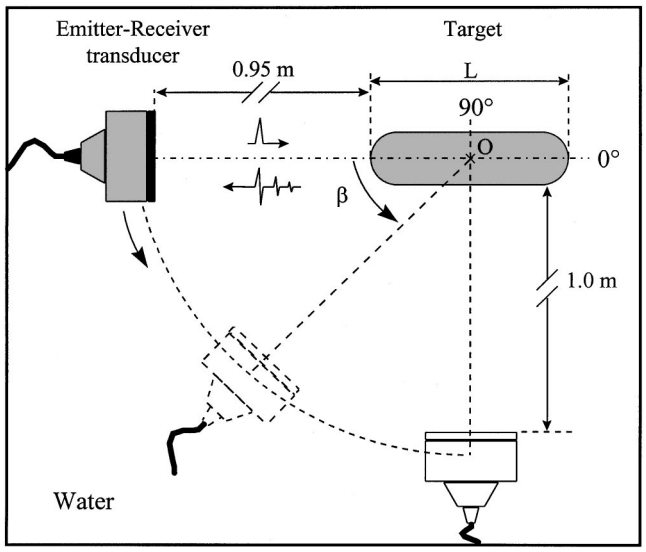


FIG. 1. Monostatic experimental setup (top view not drawn to scale).

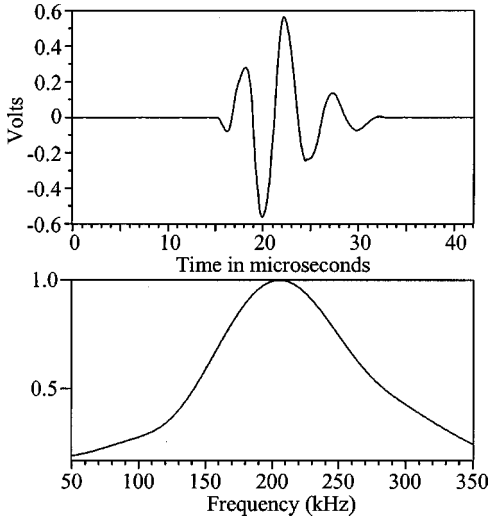


FIG. 2. Time and frequency characteristics of the transducer Panametrics, Model V3507 (specular echo and its FFT).

wave and associated modes of resonances. In Sec. IV we deal with the case of oblique incidence corresponding to the excitation of the studied shell on its cylindrical part. A geometrical model is developed to predict arrival times of different echoes and to interpret the experimental results. Finally a comparison between the acoustic response of the studied object and that of an infinite cylindrical shell is made in the frequency domain. Apart from the classical modes of resonances due to the surface helical waves observed in both cases, the main feature is the additional resonant behavior, due to the finite length of the target, present in its acoustic response which prompts a discussion.

II. EXPERIMENTAL CONDITIONS AND RESULTS

The studied object is made of stainless steel 347 of density $\rho = 7900$ kg/m^3 with the shear and longitudinal wave velocities, respectively, $C_T = 3100$ m/s and $C_L = 5790$ m/s. The dimensional characteristics are external radius $a = 15$ mm, radii ratio $b/a = 0.99$, where b is the internal radius, ratio $L/2a = 4.33$, where L is the total length of the target ($L = 130$ mm in the present case). Figure 1 presents the sketch of the experimental setup. The air-filled object, horizontally immersed in water, is excited in its meridian plane. A rectangular water tank, of 6 meters long, 4 meters wide and 3 meters deep, is equipped with a mechanism to command the displacements of the transducer around the object. The distance between the transducer and the object varies from 0.95 m to 1.0 m, depending on the transducer's position. In this study, the impulse monostatic setup[19] is used and consists in sending a short pulse to the target. The backscattered signal from the shell is received by the same transducer. The received signal, also called the echo wave form, is visualized on the oscilloscope and recorded for further signal processing and frequency analysis. The Panametrics V3507 broadband transducer, with a central frequency of 200 kHz ($k_1 a = 12.8$, where k_1 is the wave number of the incident wave in water and a the external radius) has enabled the study of the target over 60–350 kHz ($k_1 a = 4-23$). The time and frequency characteristics of the transducer are shown in Fig. 2. The diameter of its radiating surface is 50 mm. This diameter ensures the total excitation of an object for dimension $2a$ ($2a = 30$ mm). Besides in the frequency range studied and considering the distance between the transducer and the target, the aperture of the beam enables the positioning of the object in the main emission/reception lobe of the transducer whatever the incidence angle. Also, according to the pass-band of the transducer and given the physical characteristics of the shell (material and the ratio b/a) previous studies have shown that only the S_0 wave and, in oblique incidence on the cylindrical part, the T_0 wave influence the backscattering pressure.[11–15]

The echo wave form recordings made for different incidence angles starting from the axial incidence, $\beta = 0°$, up to the normal incidence, $\beta = 90°$, with a 1° angular step, are shown on the grayscale presentation in Fig. 3 with an arbi-

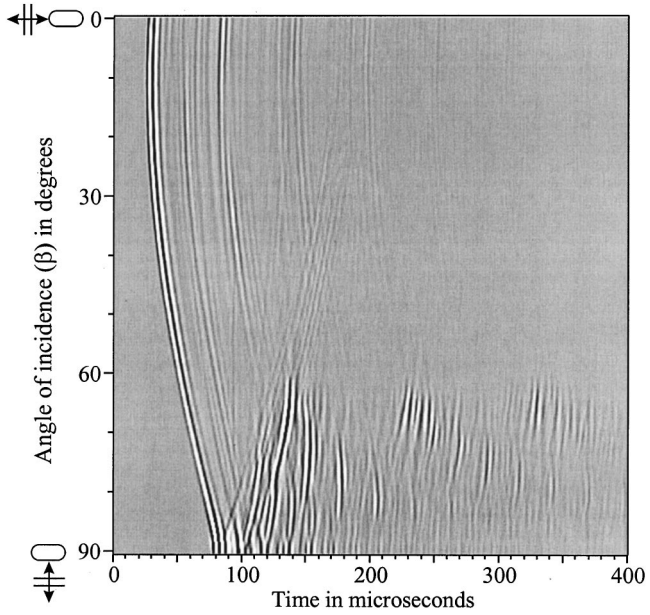


FIG. 3. Evolution of experimental echo arrival times as a function of the incidence angle β.

trary time origin. Here the magnitude of pressure is growing from black color to white. On this representation the evolution of arrival times of individual echoes can be identified as rippled bands. The first curved rippled band corresponds to the specular echo (direct reflection from the target). This band is curved in correspondence to the variation of the distance between transducer and the surface of the target for different angular positions. Besides, the transducer and the center of the target are adjusted in a line in the normal direction ($\beta=90°$) by searching the maximum amplitude of specular echo. So when the incident angle moves, the shape of specular echo becomes slightly different and shows time spreading and reduced amplitude. Two angular zones with drastically different acoustic responses, can be observed: the first zone $\beta\approx 0°-45°$, and the second one for $\beta \approx 60°-90°$. The intermediate zone ($\beta\approx 45°-60°$) is similar to the first one but with a rather low amplitude signal. Within the first zone, the surface waves are generated on the hemispherical part of the shell. Throughout this angular window, the time response of the target is not much different from that recorded in the axial position. In the second zone, the generation of the surface waves takes place on the cylindrical part of the shell. From the beginning of this zone to the end, the presence of different periodic series of echoes can be observed. In order to provide a physical interpretation of these series we will introduce a geometrical model. The two distinct angular zones are also analyzed, separately, in both time and frequency domains.

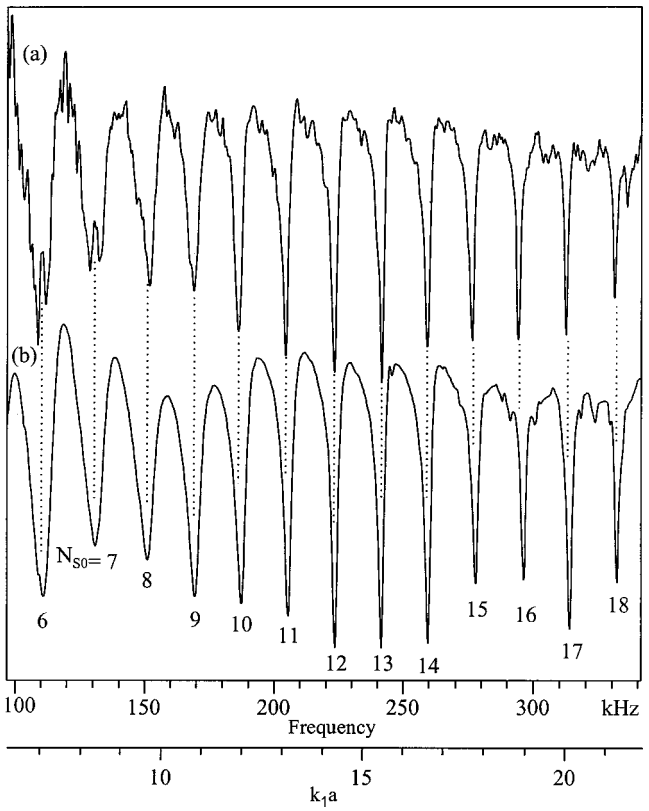


FIG. 4. Experimental backscattering spectrum (a) normalized by the FFT of the specular echo presented in Fig. 2, and calculated backscattering spectrum (b), both at the incidence angle $\beta=0°$ (axial position).

III. EXCITATION ON THE HEMISPHERICAL CAP

In this section, the studied angular zone extends from $\beta=0°$ to $\beta\approx 45°$ (Fig. 3). In the considered time window it is observed that the echo wave forms of the target are identical at that obtained for $\beta=0°$ (axial position). They do not vary significantly with the incidence angle. Therefore, only the study at one incidence angle, $\beta=0°$, in our case, is achieved. Thus numerical results obtained for the axisymmetric case can be compared not only to the recordings in the axial position, but to all angular zones of excitation on the hemispherical cap.

A. Spectral analysis

In the numerical procedure the material characteristics used are those of the stainless steel with Young's modulus $E=197.39$ GPa, Poisson's ratio $\nu=0.3$ and density $\rho=7900$ kg/m^3. Water is characterized by sound velocity $C_1=1470$ m/s and density $\rho_1=1000$ kg/m^3. The structure of the target is modeled using axi-symmetrical shell elements. The structural vibration modes are first obtained by the Finite Element Method code,[20] and are then used in the Boundary Element Method solution scheme.[21] Hence, the acoustic pressure scattered from the structure, subject to a plane wave excitation in the fluid, is determined. The modal approach chosen in this numerical study enables the reduction of both the size of the problem and the calculation time. The experimental acoustic response of the target is presented in Fig. 4(a) along with the calculated one in Fig. 4(b) in the frequency domain. The spectrum of Fig. 4(a) is obtained from the Fast Fourier Transform (FFT) of the echo wave form in axial incidence (normalized by the FFT of the specular echo shown in Fig. 2). From numerical results in Fig. 4(b), calculated modes of resonances due to the S_0 wave are indicated by N_{S0} and are in good agreement with those from the experiment [Fig. 4(a)]. These axial modes of vibration have been identified by visualizing the displacement shapes of the shell in fluid. According to these results, the object used in the experiment can be considered to be without significant internal inhomogeneities.[22] From each identified mode of resonance N_{S0} at frequency f^*, the corresponding average phase velocity $C^{\mathrm{ph}}_{S0,\mathrm{target}}$ along the meridian circumference can be obtained by applying an approximate phase matching relation:

$$C^{\mathrm{ph}}_{S0,\mathrm{target}}=\left(\frac{2\pi a+2(L-2a)}{N_{S_0}+1/2}\right)f^*. \tag{1}$$

In this relation only the phase velocity of the S_0 wave on the entire target is considered contrary to Ref. 3 where phase velocities on hemispherical and cylindrical parts are distinguished. Figure 5 shows the dispersion curves for the average phase velocity of the S_0 wave obtained from the relation (1) for both the numerical model and the experimentally studied object. It can be seen that in high frequencies the dispersion curves of the target approach the theoretical curves for a cylindrical (in axial incidence[3]) and spherical shell.[8] For low frequencies, dispersion curves of the target tend toward that of a spherical shell. All shells have the same ratio, $b/a=0.99$. Such behavior is in accordance with the

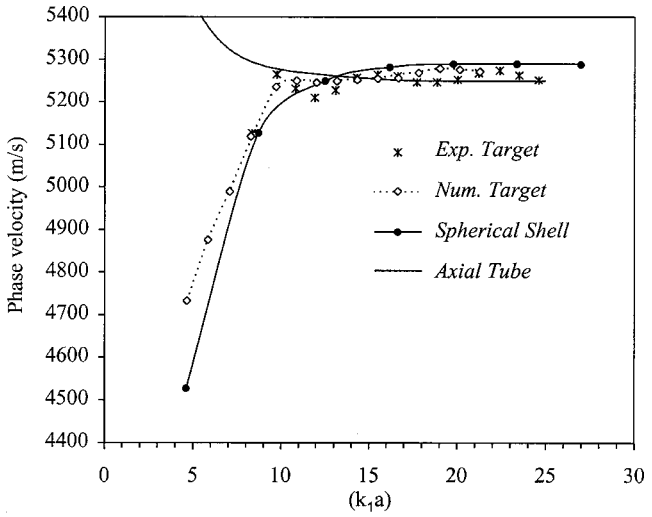


FIG. 5. Dispersion curves of phase velocities of the S_0 wave for the object studied experimentally (*) and numerically ($\diamond$) at the incidence angle $\beta = 0°$, for a spherical shell (calculations) (●) and for a cylindrical shell studied in the axial direction (calculations) (—).

fact that the influence of curvature is more significant in the lower frequency region.

B. Time domain

The experimental result at $\beta = 0°$ of Fig. 6(a) is compared with the Inverse Fast Fourier Transform (IFFT) of the calculated pressure spectrum presented in Fig. 4(b), giving the calculated echo wave form [Fig. 6(b)]. Figure 6(a) is obtained with transducer optimal settings in the axial direction whereas Fig. 3 has been recorded with transducer optimal settings in the normal direction. In this comparison, the

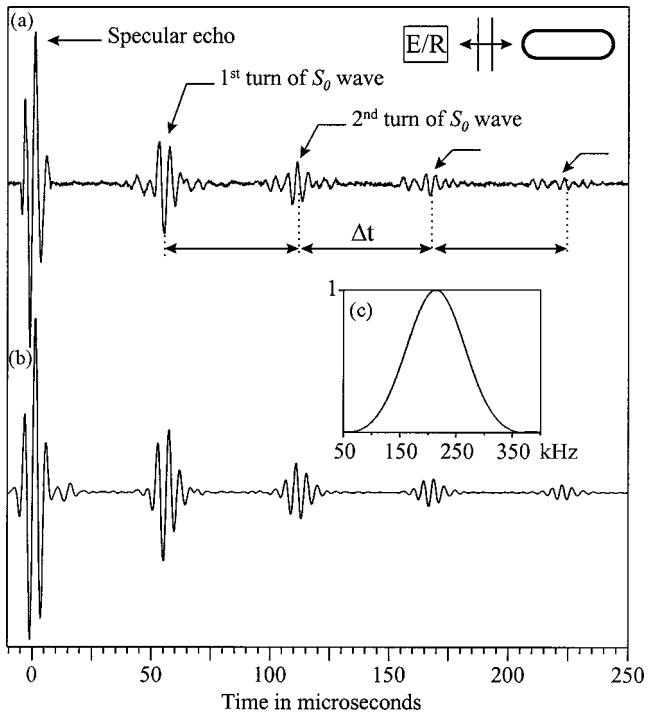


FIG. 6. Experimental (a) and calculated (b) echo wave forms obtained at the incidence angle $\beta = 0°$ (axial position), and Gaussian function centered at the frequency $f = 200$ kHz (c).

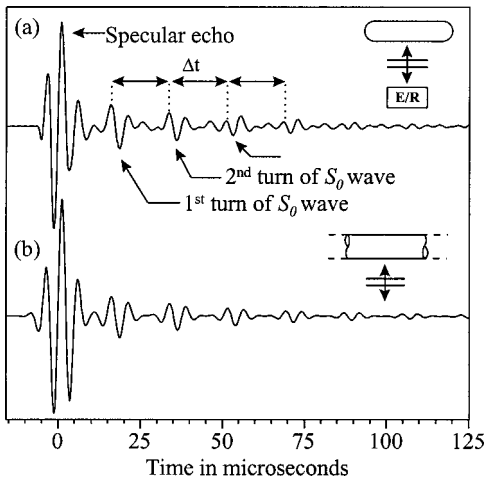


FIG. 7. Experimental echo wave form from the studied object (a) and calculated echo wave form from an infinite cylindrical shell (b), both at the incidence angle $\beta = 90°$ (normal position).

time origin is taken to the position of the specular echo. Before carrying out the IFFT, in order to show a more comprehensive comparison, the calculated spectrum has been corrected by a Gaussian function [Fig. 6(c)] to simulate the pass-band of the transducer. With the exception of the specular echo, all other echoes in Figs. 6(a)–(b) are due to the S_0 wave, generated and re-emitted on the spherical part at a characteristic angle $\theta_{S0} = \sin^{-1}(C_1/C^{ph}_{S0,sph})$ where $C^{ph}_{S0,sph}$ is phase velocity of the S_0 wave on a spherical shell. The time interval between two successive echoes, representing a single turn travel time, is measured on both experimental and calculated echo wave forms and is found to be close to 55.5 μs. This type of time response, obtained and identified in axial incidence, is present as long as the hemispherical part of the shell is excited (that is over $\beta \approx 0° - 45°$). During the measurement, the beam axis of the transducer is still lined with the target center (point O in Fig. 1). Indeed, the amplitude of signals decreases with the increase of the angle of incidence.

IV. CYLINDRICAL EXCITATION

In this section, the second angular zone observed in Fig. 3 ($60° \leqslant \beta \leqslant 90°$) is now analyzed both in time and frequency domains. In the first part of the section we consider the particular case of normal incidence ($\beta = 90°$). In the second part, the case of oblique incidence on the cylindrical part of the shell is examined for $60° \leqslant \beta < 90°$. In each case, a comparison with results from an infinite cylindrical shell is provided.

A. Normal incidence

The experimental time response at normal incidence on the cylindrical part ($\beta = 90°$) is shown in Fig. 7(a). This response is compared with a calculated echo wave form [Fig. 7(b)] obtained from the IFFT of the form function of an infinite cylindrical shell.[11] Only echoes due to the circumferential S_0 wave are present on both time responses. At this particular incidence, the re-emitted signal by hemispherical endcaps is not detected by the transducer. Except for the specular echo, each echo represents one turn of the propagat-

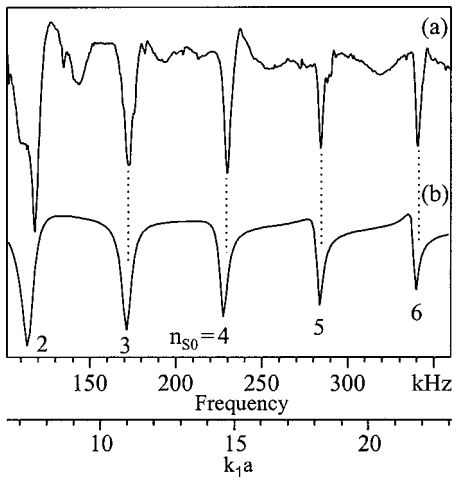


FIG. 8. A comparison of experimental backscattering spectrum from the target (a) and of the calculated backscattering spectrum from an infinite cylindrical shell (b), both at the incidence angle $\beta=90°$. Identified modes due to the circumferential S_0 wave $n_{S0}=2-6$.

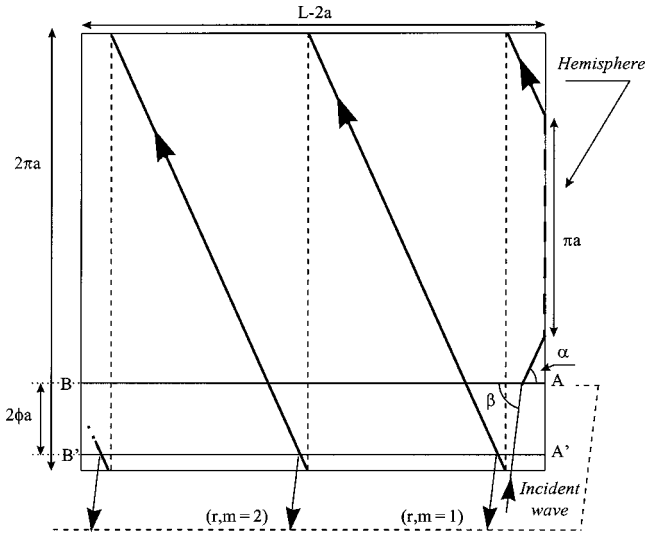


FIG. 9. Traveled paths of echoes labeled (r,m) on the unrolled cylindrical part of the target with one crossing to the reference hemisphere.

ing wave around the cylindrical cross section. Concerning the frequency domain Fig. 8 shows a good agreement between the experimental [Fig. 8(a)] and theoretical [Fig. 8(b)] backscattering spectra. The experimental spectrum [Fig. 8(a)] is normalized by the FFT of the specular echo. In the studied frequency range ($k_1a=7-23$) only modes of resonances due to the S_0 wave are observed and identified ($n_{S0}=2$ to 6). With regard to these results, when the target is excited at normal incidence it behaves exactly like an infinite cylindrical shell under the same type of excitation.

B. Oblique incidence

In the study of oblique incidence on the cylindrical part, a geometrical model is developed in order to analyze the time domain response. The study of the evolution of the echo arrival times in oblique incidence on the cylindrical part allows the identification of echo series, and therefore the interpretation of the experimental recordings. A final analysis concerns a comparison between experimental and calculated resonance spectra obtained, respectively, from our finite object and from an infinite cylindrical shell with the same shell thickness. Not only does this comparison highlight similar features, but it also confirms differences due to the finite length of the target.

1. Time domain analysis

The present theoretical model uses the formula of travel distances[18] adapted to the geometrical form of the studied target in order to predict echo arrival times. Included in the prediction of echo arrival times are the wave types with their respective generation points on the shell and their angular limits of observation.

The two-dimensional (2D) geometrical description of the model depicting the unrolled cylindrical part of the target is given in Fig. 9 for an angle of incidence β of a plane wave. On the cylindrical surface, the paths of surface waves are helical lines starting from a generating line AB for a chosen direction (or $A'B'$ for an opposite direction). When these helical lines reach the end of the cylindrical part their paths consist of semicircles on the hemisphere. The transition of the paths between the cylindrical and hemispherical parts is achieved in the same direction of propagation and without distortion due to geometrical change. Then the paths return to the cylindrical section. After an odd number r of crossings of the hemispherical endcaps, the helical wave path crosses the generating lines $A'B'$ (AB) m times, and at each crossing it radiates part of its energy toward the initial direction (monostatic scheme). As a result, patterns corresponding to odd crossings r and composed of m echoes each, can be observed in the backscattered pressure field. Thus the couple (r,m) enables us to identify the helical wave echoes.

With the radii ratio b/a of the target and the studied frequency domain only two types of waves contribute significantly to the scattering: the compressional wave $l=S_0$ and the shear wave $l=T_0$.[15] Other authors[11,18] call these waves, respectively, the Whispering-Gallery wave ($l=2$) and the guided wave ($p=1$) where (l) and (p) define the order of appearance of the resonance for circumferential and guided waves as the frequency increases. The S_0 wave is generated from the normal direction $\beta=90°$ (purely circumferential wave) to the limit angle of observation $\beta_S=90-\sin^{-1}(C_1/C_L')\approx 74°$ where C_L' is the velocity of the quasi-longitudinal wave in the plate.[23] The T_0 wave is observable from $\beta\approx 88°$ (within the threshold of oblique incidence) to the limit angle of observation $\beta_T=90-\sin^{-1}(C_1/C_T)\approx 62°$. The echo wave form recordings of Fig. 10(a) are extracted from Fig. 3 for the incidence angle range from $\beta=60°$ to $\beta=90°$. Over the total signal duration and for angles $\beta\approx 62°-88°$ three series of echoes due to three different crossings $r=1$, 3, 5 of the T_0 wave by hemispheres are observable. Besides being successive these series overlap each other. The echoes labeled (r,m) due to the T_0 and S_0 waves are indicated in Fig. 10(b). This figure shows the evolution of all these echoes as a function of the incidence angle obtained from calculations using arrival time formulas which are developed in the next paragraph. The evolution of the calculated arrival times matches with that obtained experimentally.

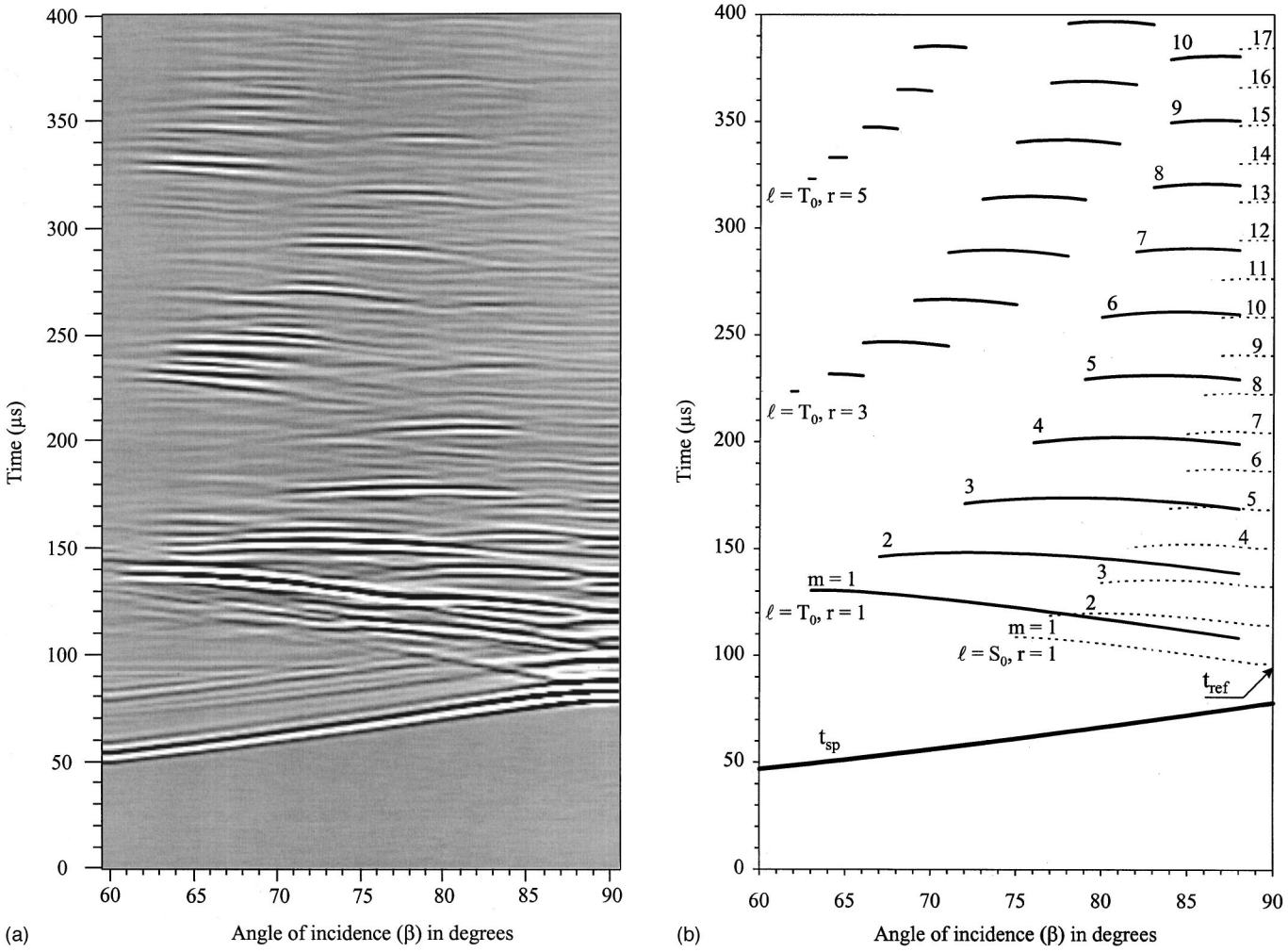


FIG. 10. (a) Evolution of experimental echo arrival times as a function of the incidence angle β. (b) Evolution of calculated echo arrival times as a function of the incidence angle β.

Relations between the angles θ_l, α, β and ϕ used in the model, are[17]

$$\alpha = \cos^{-1}(\cos\beta/\sin\theta_l), \quad 0° < \alpha < 90°, \tag{2}$$

$$\phi = \cos^{-1}(\cos\theta_l/\sin\beta), \quad 0° < \phi < \theta_l, \tag{3}$$

where $\theta_l = \sin^{-1}(C_1/C_l)$ is the characteristic angle for the l type wave, β is the angle of the incident direction with the target axis, α is the helical angle of wave, and ϕ is the half opening angle between generating lines AB and A′B′ (Fig. 9).

The travel distance along the surface of the cylindrical part can be written as

$$D_{\text{cyl}}(r,m) = 2(m\pi - r\pi/2 - \phi)a/\sin\alpha, \tag{4}$$

the travel distance along the surface of the spherical part is

$$D_{\text{sph}}(r) = r\pi a, \tag{5}$$

and the travel distance in water:

$$D_{\text{wat}}(r,m) = 2a(1-\cos\phi)\sin\beta + [(L-2a)(r+1) - 2\,\Delta h(r,m)]\cos\beta - (L-2a)\cos\beta - 2a\sin\beta, \tag{6}$$

with

$$2\Delta h(r,m) = 2(m\pi - \phi)a/\tan\alpha - r\pi a/\tan\alpha, \tag{7}$$

which represents the geometrical distance along the cylinder between the point of wave generation and the point of observation of helical waves. This distance is then adjusted by the path traveled on the hemispheres. The relation (7) is determined under the following conditions:

$$\Delta h(1,1) \leqslant \Delta h(r,m) \leqslant L - 2a, \quad \text{for } r = 1 \tag{8}$$

and

$$(r-1)(L-2a) \leqslant 2\Delta h(r,m) \leqslant (r+1)(L-2a), \quad \text{for } r > 1. \tag{9}$$

These two conditions enable the determination of the minimum and maximum numbers of observable echoes ($m_{\min}$ and $m_{\max}$) for the shortest and the longest cylindrical paths. Thus echo arrival times of the specular echo is

$$t_{\text{sp}}(\beta) = t_{\text{ref}}(\beta = \pi/2) - ((L-2a)\cos\beta + 2a\sin\beta)/C_1 + \Delta t_{\text{ref}}, \tag{10}$$

where $t_{\text{ref}}(\beta = \pi/2)$ represents the arrival time of the first S_0 wave echo (just after the specular echo) at normal incidence.

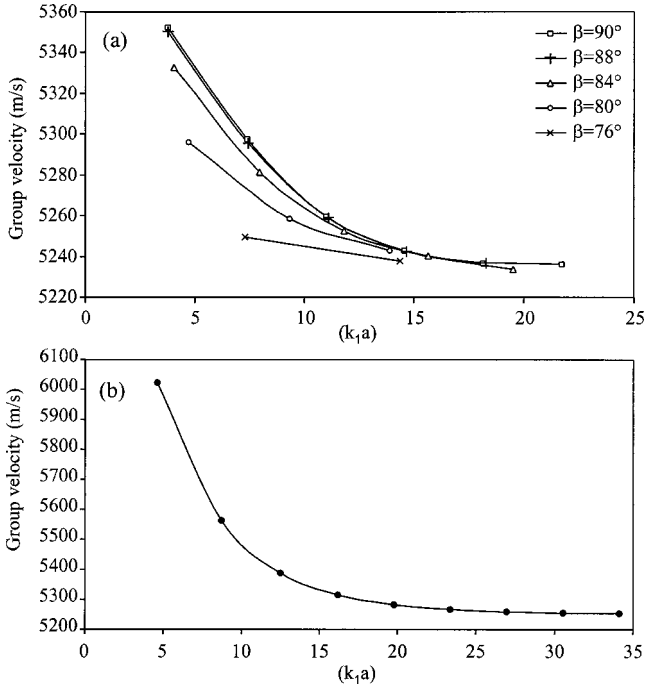


FIG. 11. Theoretical dispersion curves of group velocities of the S_0 wave, respectively, from an infinite cylindrical shell in oblique incidence (a) and from a spherical shell, both in water (b).

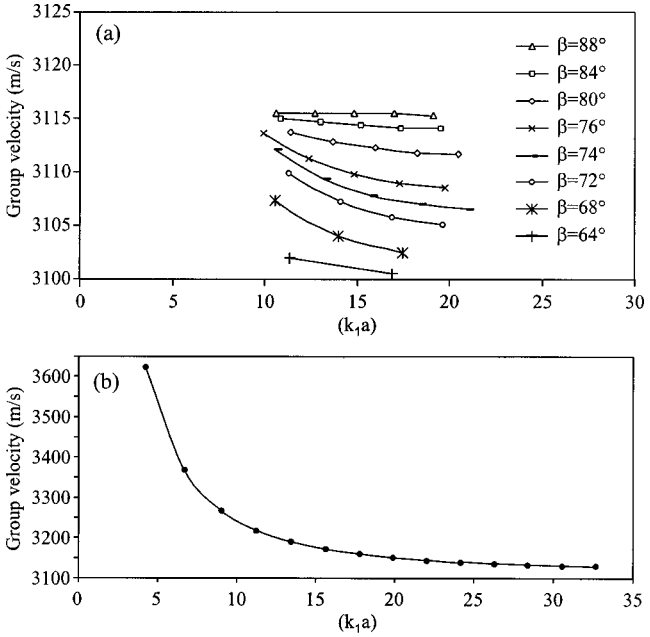


FIG. 12. Theoretical dispersion curves of group velocities of the T_0 wave, respectively, from an infinite cylindrical shell in oblique incidence immersed in water (a) and from a spherical shell in vacuum (b).

The choice of this reference echo is arbitrary and facilitates a comparison with the experimental echo arrival times. Time shift $\Delta t_{\text{ref}}=2a[\cos\theta_{S0}/C_1-(\pi-\theta_{S0})/C^g_{S0,\text{cyl}}]$ is due to the particular choice of reference time.

For the echoes of each crossing, the arrival times can be expressed as

$$t(r,m)=D_{\text{cyl}}(r,m)/C^g_{l,\text{cyl}}+D_{\text{sph}}(r)/C^g_{l,\text{sph}}+D_{\text{wat}}(r,m)/C_1+t_{\text{ref}}(\beta=\pi/2)+\Delta t_{\text{ref}}, \quad (11)$$

where $C^g_{l,\text{cyl}}$ and $C^g_{l,\text{sph}}$ are the group velocities, respectively, along the cylindrical and on the spherical surfaces, for the considered l type wave. These velocities are calculated for spherical[8] and infinite cylindrical shells.[11] Calculations are achieved with shells immersed in water, except for the T_0 modes on a spherical shell which are determined in vacuum.[24] Although all experiments are carried out in water, theoretical results obtained in vacuum can be used in comparisons because the water-shell coupling is known to be weak for this wave. Figures 11 and 12 show the theoretical dispersion curves of the S_0 and T_0 waves, respectively, for the cylindrical and then for the spherical shells. According to the results from these figures within the frequency range of the study, the group velocities can be taken as

$$C^g_{S0,\text{sph}}=C^g_{S0,\text{cyl}}\approx 5230 \text{ m/s}, \quad (12)$$

$$C^g_{T0,\text{sph}}=C^g_{T0,\text{cyl}}\approx 3110 \text{ m/s}. \quad (13)$$

Although, in the considered frequency domain, the group velocities of these two wave types are slightly dispersive, these velocities are considered to be constant in the calculation of the time response. They correspond to asymptotic values for large values k_1a and for a different incidence angle.

Furthermore concerning the T_0 wave on the spherical shell,[24] owing to the difficulty in generating it directly on the spherical surface it has never been observed experimentally. However for the studied object this wave type can propagate on the hemispherical shells due to the obliquely excitation of the cylindrical part.

2. Frequency domain analysis

In this paragraph a comparison between experimental resonance spectra for the cylindrical shell bounded by hemispheres and calculated resonance spectra for an infinite cylindrical shell is presented. Figures 13(a) and (b) illustrate, respectively, the evolution of experimental and calculated resonances as a function of the incidence angle. For the angular window $61°\leqslant\beta\leqslant 90°$ the FFT of experimental time responses [Fig. 10(a)] is calculated from $t=210$ μs (corresponding approximately to the beginning of the crossing $r=3$) to the end of recordings $t=1000$ μs. The choice of $t=210$ μs has been made in order to obtain the free response of the target. The resonance spectra are obtained in the dimensionless frequency range of $3\leqslant k_1a\leqslant 22$ around the central frequency of the transducer $k_1a=12.8$ ($f=200$ kHz). Concerning the infinite cylindrical shell in the far-field, the form functions are calculated for the same angular and frequency windows, but the sound pressure spectra are calculated in the Snell–Descartes position of the receiver. From these calculated results the IFFT can be also obtained, and then the FFT of these last time domain results excluding the specular echo, gives resonance spectra.

The most significant feature in Figs. 13(a)–(b) is the matching resonant behaviors of the studied object and infinite cylindrical shells. Resonances due to S_0 and T_0 helical wave types appear in the experimental spectra (top part of the figure) as scratched black curves. In the calculated spec-

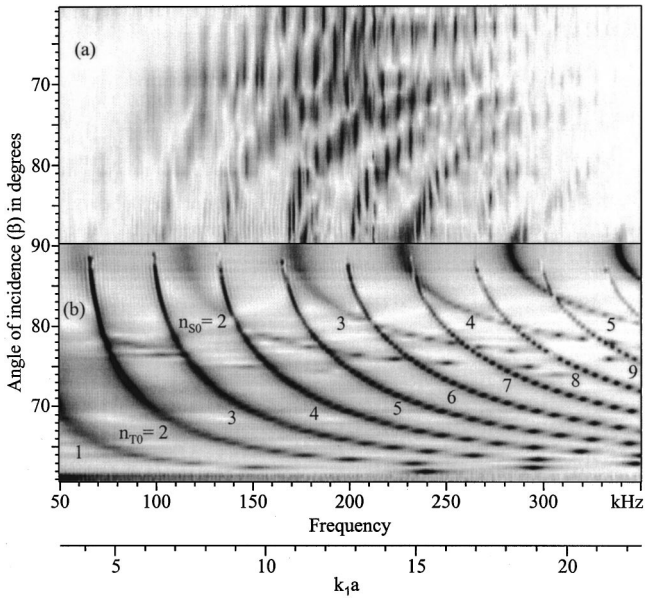


FIG. 13. Evolution of experimental resonance spectra from the excitation on the cylindrical part of the studied object (a) and of calculated resonance spectra from an infinite cylindrical shell (b), both as a function of the incidence angle β.

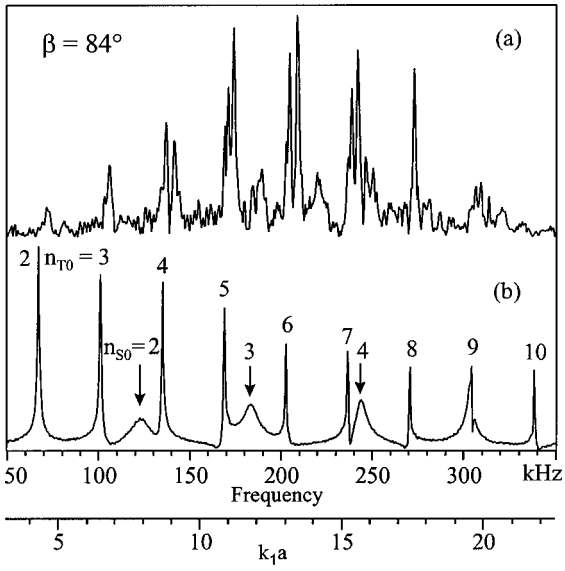


FIG. 14. Experimental resonance spectrum obtained from the studied object on the cylindrical part (a) and calculated resonance spectrum issued from an infinite cylindrical shell (b), both at the incidence angle $\beta=84°$. Calculated modes of resonance due to the helical S_0 wave ($n_{S0}=2-4$) and to the helical T_0 wave ($n_{T0}=2-10$).

tra of the infinite shell (bottom part of the figure) the black curves must be regarded as continuous, discrete appearance of these lines is due only to the angular calculation step. All modes due to the S_0 wave converge to the limit angle $\beta \approx 74°$ (clearly observable on calculated spectra), and those due to the T_0 wave converge to the limit angle, $\beta \approx 61°$. It can be noted that the tube resonance modes (n_{T0}, n_{S0}) for an infinite cylindrical shell in the Snell–Descartes position have the same frequency-angle evolution as those of a finite cylindrical shell bounded by hemispheres in the backscattering position. The main difference observed between the experimental and calculated spectra in Figs. 13(a)–(b) are the scratches present for the studied object. In fact, the scratches are due to the FFT of the echo series issued from different hemisphere crossings r (and consequently ''knowing'' about the length of the object), which overlaps with each other in time, as can be seen in Figs. 10(a) and (b). Indeed, the FFT of only the first series of echoes of the crossing $r=1$ gives quite smooth curves without scratches. Figure 14 shows this difference between the studied object [Fig. 14(a)] and infinite cylindrical shell [Fig. 14(b)], both at the incidence angle $\beta = 84°$. Note that the experimental frequency positions of the resonances due to the T_0 wave match with those calculated for an infinite tube. In Fig. 14(a) several peaks of resonance around calculated frequency positions are present. The presence of these resonances is related to the formation of closed paths on the studied shell. To highlight the influence of the shell length, the experimental resonance spectrum at the incidence angle $\beta=61°$ is presented in Fig. 15. This angular position is chosen because only the first echoes of each crossing r are present and consequently both time domain and frequency domain results are easier to interpret. In particular, modal density at this angle permits the separation of resonances, so that identification by using the phase matching condition[17] can be applied:

$$\left(\frac{2(L-2a)\sin\theta_l}{C^{\mathrm{ph}}_{l,\mathrm{cyl}}\cos\beta}+\frac{2\pi a}{C^{\mathrm{ph}}_{l,\mathrm{sph}}}\right)f^*=\tilde{N}+\frac{1}{2},\quad \pi/2-\theta_l\leqslant\beta<\pi/2. \tag{14}$$

In this relation, for a particular angle of incidence β and considering an l helical surface wave propagating with the phase velocities $C^{\mathrm{ph}}_{l,\mathrm{cyl}}$ and $C^{\mathrm{ph}}_{l,\mathrm{sph}}$ on cylindrical and spherical parts, respectively, two helical lines along the surface of the cylindrical part, and two semicircles on the hemispheres are considered to form a closed path corresponding to a mode of resonance $\tilde{N}$ at the resonance frequency f^*. The dispersion curves of the respective phase velocities $C^{\mathrm{ph}}_{l,\mathrm{cyl}}$ and $C^{\mathrm{ph}}_{l,\mathrm{sph}}$ are estimated for each l wave from studies based on the vibration modes of an infinite cylindrical shell[11] and a spherical shell.[8] Several helical modes $\tilde{N}_{T0}$ due to the T_0 wave are indicated

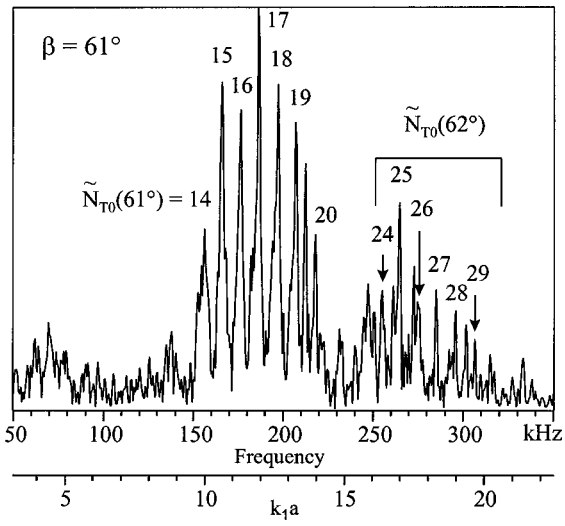


FIG. 15. Experimental resonance spectrum at the incidence angle $\beta=61°$ obtained with the excitation of the studied object on the cylindrical part.

TABLE I. Resonance modes of the T_0 wave for an extra thin tube bounded by hemispherical shells at the incidence angle $\beta=61°$.

Incidence angle $\beta=61°$						
Experimental frequency		Phase velocity Spherical shell	Cylindrical shell	Mode $\tilde{N}_{T0}$ Phase matching		Calculated frequency
kHz	k_1a	$C_{\text{sph}}^{\text{ph}}(T_0)$ m/s	$C_{\text{cyl}}^{\text{ph}}(T_0)$ m/s	Calculated	Integer	kHz
156.3	10.0	2970	3100	14.32	14	152.9
166.0	10.6	2980	3100	15.23	15	163.6
176.5	11.3	2990	3100	16.20	16	174.4
186.3	11.9	3010	3100	17.09	17	185.3
197.3	12.6	3020	3100	18.10	18	196.1
207.0	13.3	3030	3100	19.00	19	206.9
218.0	14.0	3040	3100	20.01	20	217.8

in Fig. 15. It is clearly observed that these modes are centered around the central frequency of the transducer. They correspond to the scratches present in Fig. 13(a) at $\beta=61°$ with the frequency gap between successive resonance scratches approximately equal to $\Delta f\approx 10$ kHz. Other peaks of resonances situated in higher frequencies are observable. These peaks result from modes identified at $\beta=62°$ and are due to the effect of transducer aperture. Tables I and II summarize these experimental results.

V. CONCLUSION

Acoustic response of an obliquely excited thin cylindrical shell bounded by hemispherical endcaps is described and interpreted. The acoustic response of the studied object is drastically different depending on excitation produced on hemispherical or cylindrical parts of the shell.

In the case of excitation of the hemispherical part of the shell, the response remains very similar at all incidence angles. In fact it is not much different from the axisymmetric excitation case and corresponding numerical results can be helpful for geometrical interpretations. Considering the thickness of the shell and frequency range explored, only S_0 wave circumnavigating around the great circumference can be generated and observed.

In the case of excitation of the cylindrical part of the shell, the S_0 and T_0 helical waves are generated. The different series of echoes corresponding to the re-emission of a propagating helical wave are observed and identified. In the frequency domain the resonances series can be identified by an application of the phase matching condition. These resonances on the closed path due to the finite length of the object can be observed only in the case of angular and frequency coincidences with one of the infinite tube modes.

TABLE II. Resonance modes of the T_0 wave for an extra thin tube bounded by hemispherical shells at the incidence angle $\beta=62°$.

Incidence angle $\beta=62°$						
Experimental frequency		Phase velocity Spherical shell	Cylindrical shell	Mode $\tilde{N}_{T0}$ Phase matching		Calculated frequency
kHz	k_1a	$C_{\text{sph}}^{\text{ph}}(T_0)$ m/s	$C_{\text{cyl}}^{\text{ph}}(T_0)$ m/s	Calculated	Integer	kHz
255.3	16.4	3060	3100	23.99	24	255.3
265.1	17.0	3065	3100	24.92	25	265.8
275.5	17.7	3070	3100	25.91	26	276.4
285.2	18.3	3070	3100	26.84	27	286.8
296.2	19.0	3070	3100	27.89	28	297.3
306.6	19.7	3070	3100	28.89	29	307.7

ACKNOWLEDGMENTS

The authors would like to thank their colleague Dr. J. G. Chiumia for his useful comments on the text while in preparation and his encouragement.

[1] G. Maze, F. Lecroq, D. Décultot, J. Ripoche, S. K. Numrich, and H. Überall, ''Acoustic scattering from finite cylindrical elastic objects,'' J. Acoust. Soc. Am. **90**, 3271–3277 (1991).
[2] R. D. Miller, E. Thomas Moyer, Jr., H. Huang, and H. Überall, ''A comparison between the boundary element method and the wave superposition approach for the analysis of the scattered fields from rigid bodies and elastic shells,'' J. Acoust. Soc. Am. **89**, 2185–2196 (1991).
[3] D. Décultot, F. Lecroq, G. Maze, and J. Ripoche, ''Acoustic scattering from a cylindrical shell bounded by hemispherical endcaps. Resonance interpretation with surface waves propagating in cylindrical and spherical shells,'' J. Acoust. Soc. Am. **94**, 2916–2923 (1993).
[4] B. Dubus, A. Lavie, and N. D. Veksler, ''Frequency derivative approach for identification of the wave resonances on immersed elastic bodies,'' J. Acoust. Soc. Am. **102**, 3523–3529 (1997).
[5] A. Klauson, J. Metsaveer, N. Touraine, D. Décultot, and G. Maze, ''Sound scattering by a cylindrical shell with hemispherical endcaps,'' *Transport Noise 98*, Joint International EAA/EEAA Symposium, Tallinn, Estonia, Conference Proceedings, 1998, pp. 243–246.
[6] D. Felix, ''Numerical model for acoustic radiation of an immersed elastic structure and application to a thin cylindrical shell,'' J. Acoust. Soc. Am. **95**, 3332–3338 (1994).
[7] G. Maze, F. Léon, and J. Ripoche, ''Nature de l'onde d'interface de Scholte sur une coque cylindrique,'' Acustica **81**, 201–213 (1995).
[8] M. Talmant, H. Überall, R. D. Miller, M. F. Werby, and J. W. Dickey, ''Lamb waves and fluid-borne waves on water-loaded, air-filled thin spherical shells,'' J. Acoust. Soc. Am. **86**, 278–289 (1989).
[9] N. Touraine, J. Chiumia, Ch. Barrière, D. Décultot, G. Maze, A. Klauson, and J. Metsaveer, ''Influence of the internal solder layer and of A wave on acoustic scattering from finite cylindrical shells with hemispherical endcaps,'' *4th European Conference on Underwater Acoustics*, Rome, Italy, Conference Proceedings, 1998, pp. 897–902.
[10] H. Überall, ''Helical surface waves on cylinders and cylindrical cavities,'' Trait. Signal **2**, 381–385 (1985).
[11] F. Léon, F. Lecroq, D. Décultot, and G. Maze, ''Scattering of an obliquely incident acoustic wave by an infinite hollow cylindrical shell,'' J. Acoust. Soc. Am. **91**, 1388–1397 (1992).
[12] F. Lecroq, F. Léon, D. Décultot, and G. Maze, ''Diffusion d'une onde acoustique par un tube limité par des disques plans: études théoriques et expérimentales,'' Acustica **74**, 51–62 (1991).
[13] J. M. Conoir, P. Rembert, O. Lenoir, and J. L. Izbicki, ''Relation between surface helical waves and elastic cylinder resonances,'' J. Acoust. Soc. Am. **93**, 1300–1307 (1993).
[14] N. D. Veksler, G. Maze, J. Ripoche, and V. Porochovskii, ''Scattering of an obliquely incident plane acoustic wave by a circular cylindrical shell. Results of computations,'' Acustica-Acta Acustica **82**, 689–697 (1996).
[15] S. F. Morse, Ph. L. Marston, and G. Kaduchak, ''High-frequency backscattering enhancements by thick finite cylindrical shells in water at oblique incidence: Experiments, interpretation, and calculations,'' J. Acoust. Soc. Am. **103**, 785–794 (1998).
[16] M. L. Rumerman, ''Contribution of membrane wave reradiation to scattering from finite cylindrical steel shells in water,'' J. Acoust. Soc. Am. **93**, 55–65 (1993).
[17] X.-L. Bao and H. Überall, ''Experimental study of acoustic resonances of elastic spheres and hemispherically endcapped cylinders,'' in *Acoustic Resonance Scattering*, edited by H. Überall (Gordon and Breach, Philadelphia, 1992).
[18] X.-L. Bao, ''Echoes and helical surface waves on a finite elastic cylinder excited by sound pulses in water,'' J. Acoust. Soc. Am. **94**, 1461–1466 (1993).
[19] P. Pareige, P. Rembert, G. Maze, and J. Ripoche, ''Signature spectrale et

identification des résonances d'objets immergés par une méthode impulsionnelle numérisée (MIIR impulsionnelle),'' J. Acoust. **3**, 101–106 (1990).
[20] Engineering Mechanics Research Corporation, NISA II/DISPLAYS III, EMRC, P. O. Box 696, Troy, Michigan 48099.
[21] Numerical Integration Technologies, theoretical manual of SYSNOISE, NIT N. V., Ambachtenlaan 11a, B-3001 Leuven, Belgium.
[22] N. Touraine, A. Klauson, J. Metsaveer, D. Décultot, and G. Maze, ''Effect of internal inhomogeneities on acoustic scattering from finite cylindrical shell bounded by hemispherical endcaps,'' *Forum Acusticum 1999, 137th Meeting Acoustical Society of America, 2nd convention of the EAA and 25th German Acoustics DAGA Conference*, Berlin, Germany, 1999, ACUSTICA-acta acustica Vol. 85, Suppl. 1, Abstract, S162.
[23] The velocity of plate quasi-longitudinal wave is $C_L'=2C_T[2-(C_L^2-2C_T^2)/(C_L^2-C_T^2)]^{-1/2}$, using the parameters of Sec. II, C_L'(SS347)=5237 m/s.
[24] C. Eringen and E. S. Suhubi, *Elastodynamics: Linear Theory* (Academic, New York, 1975), Vol. 2.

Acoustic radiation due to the inelastic impact of a sphere on a rectangular plate

Philippe Troccaz,[a] Roland Woodcock,[b] and Frédéric Laville[c]
Université du Québec, Ecole de Technologie Supérieure, Département de Génie Mécanique, 1100, rue Notre-Dame Ouest, Montréal, Québec H3C 1K3, Canada

(Received 12 January 1996; revised 25 July 2000; accepted 27 July 2000)

Impact between two structures is an important source of noise in industry. However, it is not a well-known phenomenon from a theoretical point of view. The models available in the literature for impact noise often address only part of the phenomenon. In this paper, a more comprehensive model is developed. The acoustic radiation due to the inelastic impact of a sphere with a rectangular simply supported thin plate is formulated and validated experimentally. The impact force is calculated from Hertz's law, which has been extended to the case of plastic deformations, and the acoustic radiation in the time domain is obtained using Rayleigh's integral, the plate being discretized in rectangular elements. The model is used to show that the sound-radiation mechanism includes two separate sources: the initial deformation of the plate and the propagation of the bending waves. The model is also used to determine the contribution to impact noise of parameters that characterize the two impacting structures or the impact configuration such as plasticity, damping, materials, and impact velocity. © *2000 Acoustical Society of America.* [S0001-4966(00)01011-0]

PACS numbers: 43.40.Rj, 43.40.Kd [CBB]

I. INTRODUCTION

Impact between two structures is an important source of noise in industry. However, it is not a well-known phenomenon from a theoretical point of view. In vibro-acoustical problems, excitation forces can be split into two main categories; excitation forces which are independent of the structure and determined *a priori*, and those that result from a source–structure interaction and thus cannot be known independently from the structure. The forces of the first category are encountered, for instance, in the case of an acoustical excitation or some mechanical excitations such as an unbalanced rotating shaft. In the case of the impact excitation of concern in this paper, the excitation forces are of the second category. Consequently, it is necessary to consider the calculation of the impact force in addition to the calculation of the acoustic radiation.

For the impact force calculation, a first approach consists of adopting elastic deformations,[1] but this approach is an approximation which is far from that encountered in real situations. It is therefore necessary to take into account inelastic deformations.[1–5]

For the acoustic radiation calculation, the collision between two spheres,[6,7] between a sphere and a circular plate with the impact force approximated by a trigonometric function[8,9] and between a sphere and an infinite plate from a frequency point of view[10] have been presented.

The aim of the research presented in this paper was to develop a more comprehensive model including calculations for both the impact force and the acoustic radiation in the case of the inelastic impact of a sphere on a simply supported rectangular plate.

The theoretical formulation is presented in Sec. II followed by the simulation results and the experimental validation in Sec. III, and a parametric study in Sec. IV.

II. THEORETICAL FORMULATION

The acoustic radiation due to the impact is obtained in three successive calculations: first, the impact force, second, the response of the plate to the impulsive load, and, third, the acoustic pressure at any point of space.

A. Calculation of the impact force

A plate of finite dimensions a and b and thickness h is illustrated in Fig. 1. The Love–Kirchhoff model for the plate is used based on the assumption that the plate only carries bending waves. The fluid surrounding the plate is homogeneous, with a speed of sound c and a density ρ_0 much less than ρ, the density of the plate. The collision between the sphere and the plate is located at a point of coordinates x_0 and y_0 in the plate coordinate system. At the impact point, the sphere has an initial velocity V_0. The impact force is obtained by solving a set of three equations[11]

The first equation comes from Newton's law applied to the sphere motion

$$\frac{d^2}{dt^2}(u(t)+w(t))+\frac{1}{m}F(t)=0, \tag{1}$$

where m is the mass of the sphere, $u(t)$ is the displacement of the sphere relative to the plate, $w(t)$ is the displacement of the plate at the contact point, $u(t)+w(t)$ is the absolute displacement of the sphere, and $F(t)$ is the impact force.

The second equation gives the plate response

[a]Present address: EDF-CNEN 165-173, avenue Pierre Brossolette, BP 900, 92542 Montrouge cedex, France.
[b]Present address: Advanced Acoustics Group, Lear Technologies, LLC, N. American Technology Center, 21557 Telegraph Road, Southfield, MI 48034-4248.
[c]Author to whom correspondence should be addressed. Electronic mail: frederic.laville@etsmtl.ca

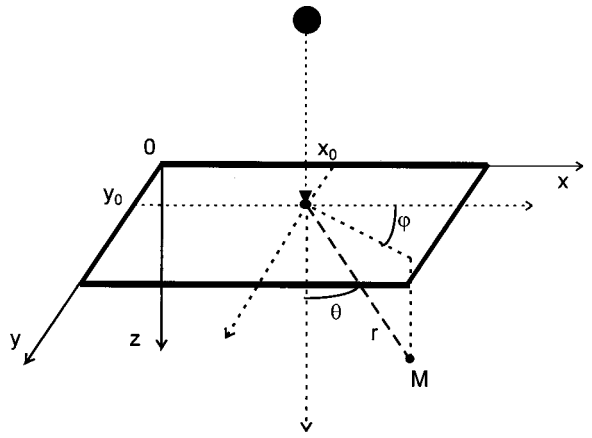


FIG. 1. Impact configuration.

$$w(t)=\int_0^t F(t)M(t-\tau)d\tau, \tag{2}$$

where $M(t)$ is the impulse response of the plate, defined by the following equation:

$$M(t)=\frac{4}{\rho hab}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\sin^2\left(\frac{n\pi x_0}{a}\right)\sin^2\left(\frac{m\pi y_0}{b}\right)\times\frac{\sin(\theta_{n,m}t)}{\theta_{n,m}}e^{-\delta_{n,m}t}$$
$$\theta_{n,m}=\omega_{n,m}\sqrt{1-\frac{\xi^2}{4}}\quad\text{and}\quad\delta_{n,m}=\frac{\xi}{2}\omega_{n,m}. \tag{3}$$

ξ is the damping coefficient of the plate and $\omega_{n,m}$ the natural frequency for the (n,m)th mode, given by

$$\omega_{n,m}=\sqrt{\frac{D}{\rho h}}\left[\left(\frac{n\pi}{a}\right)^2+\left(\frac{m\pi}{b}\right)^2\right]\quad\text{and}\quad D=\frac{Eh^3}{12(1-\nu^2)}, \tag{4}$$

where E and ν are, respectively, the Young's modulus and the Poisson's coefficient of the plate.

The third equation describes the relation between the impact force and the sphere displacement. If the impact is assumed to be elastic, this relation is given by Hertz's law[1]

$$F(t)=Ku(t)^{3/2} \tag{5}$$

where K is the Hertz's constant, which depends on parameters of both the sphere and the plate

$$K=\frac{4}{3}\sqrt{R}\left[\frac{1-\nu_1^2}{E_1}+\frac{1-\nu^2}{E}\right]^{-1}, \tag{6}$$

where R is the sphere radius, E_1 and ν_1 are, respectively, the Young's modulus and the Poisson's coefficient of the sphere.

As Hertz's law can only be applied for elastic impacts and as plastic deformations often occur in impacts even at very low speeds, a model extending Hertz's law to take into account plastic deformations has been developed. The contact period is divided into three phases.

(i) The first phase is the elastic deformation: $F(t)=Ku(t)^{3/2}$.

(ii) The second phase is the purely plastic deformation, as soon as the constraint on the deformation surface reaches P_0, the dynamic yield strength of the deforming material. The value of P_0 for aluminum is given in a table in Ref. 1 and, for other materials, it can be obtained from empirical relations given in Ref. 4. The expression of the force is then $F(t)=2\pi RP_0u(t)$. This phase ends when the force reaches its maximum value $F_{\max}$.

(iii) The third phase is the elastic deformation during the rebound, with $F(t)=F_{\max}([u(t)-u_r]/[u_m-u_r])^{3/2}$, where u_r is the residual displacement (plastic) and u_m is the maximum displacement (elastic and plastic).

This calculation was obtained by combining several theoretical models. The model for the first phase is similar to the one in Ref. 2, where the impacting sphere deforms elastically and plastic strains are neglected. The model for the second phase is taken from Ref. 3, where only two phases are considered (plastic deformation and elastic recovery), the elastic regime before the onset of plasticity being neglected. The model for the third phase comes from Ref. 5, where the model is based on experimental data for the indentation process, the recovery process during the rebound being Hertzian.

This approach is based on three simplifying hypotheses. First, it is assumed that the three phases occur separately, one after the other. Second, the dynamic yield strength P_0, which depends on the impact velocity, is assumed to be constant in the range of velocities under concern in the present paper. Finally, only one of the two colliding bodies deforms plastically.

This calculation of the force has not been validated itself. However, the numerical results of the acoustic pressure calculated using this computation of the force present a good agreement with the corresponding measurements, as will be shown in the next part of the present paper. The system is solved numerically using the Newton–Raphson method, the initial conditions being that the plate is at rest and the sphere has an initial velocity V_0.

B. Calculation of the plate response

The general solution of the equation of motion of a finite rectangular plate subjected to a point force $F(t)$ is given by[1,9]

$$w(x,y,t)=\frac{4}{\rho hab}\sum_i\sum_j\frac{H_{ij}(x,y)H_{ij}(x_0,y_0)}{\omega_{ij}^*}\times\int_0^t F(t)e^{-\xi\omega_{ij}(t-\tau)}\sin[\omega_{ij}^*(t-\tau)]d\tau, \tag{7}$$

where

$$H_{ij}(x,y)=\sin\left(\frac{i\pi x}{a}\right)\sin\left(\frac{j\pi y}{b}\right)$$

are the eigenmodes of the plate,

$$\omega_{ij}=\sqrt{\frac{D}{\rho h}}\left[\left(\frac{i\pi}{a}\right)^2+\left(\frac{j\pi}{b}\right)^2\right]$$

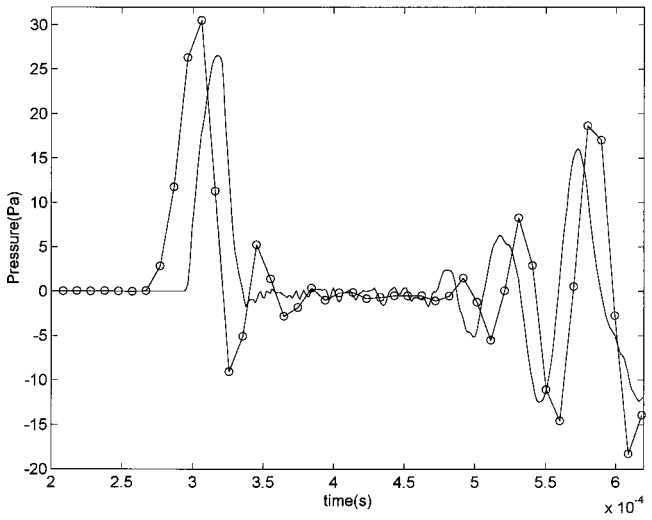


FIG. 2. Acoustic pressure for $\theta=\varphi=0°$, $r=10$ cm, $V_0=1.4$ m/s, standard aluminum plate, steel sphere, numerical (solid) and experimental (dashed).

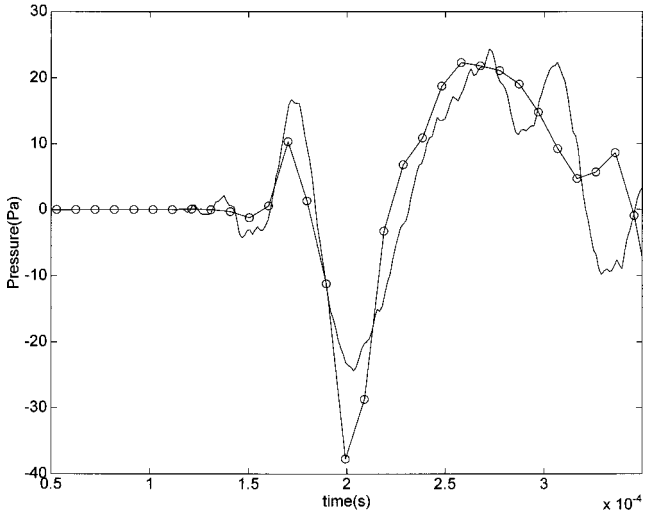


FIG. 3. Acoustic pressure for $\theta=73°$, $\varphi=0°$, $r=10.6$ cm, $V_0=1.4$ m/s, standard aluminum plate, steel sphere, numerical (solid) and experimental (dashed).

are the natural frequencies of the plate, and $\omega_{ij}^{*}=\omega_{ij}\sqrt{1-\xi^2}$.

C. Calculation of the acoustic pressure

The radiated pressure at a point $M(r,\theta,\varphi)$ is calculated using Rayleigh's formula

$$p(r,\theta,\varphi,t)=\frac{\rho_0}{2\pi}\int_S \frac{\ddot{w}[x,y,t-(d/c)]}{d}\,dS, \tag{8}$$

where $d=\sqrt{(x-x_M)^2+(y-y_M)^2+z_M^2}$,

$x_M=x_0+r\sin\theta\cos\varphi$, $\quad y_M=y_0+r\sin\theta\sin\varphi$,

$z_M=r\cos\theta$.

The calculation of this integral is performed numerically by discretizing the plate into rectangular surface elements and then summing the contribution of each element considered as a point source with the appropriate time delay d/c.

III. SIMULATION RESULTS INTERPRETATION AND EXPERIMENTAL VALIDATION

The simulation results are presented in the case of the impact between a 7.8-mm-diameter steel sphere and a 405 $\times$325$\times$3-mm aluminum plate. The sphere initial velocity is 1.40 m/s and the impact occurs at the center of the plate. Numerical calculations have been carried out using a 10 000-point spatial discretization and taking into account the first 100 modes of the plate. Two microphone positions are considered, one near the normal to the plate at the impact point (small θ) and the other one off the normal (large θ).

In both cases the simulation results are compared to experimental results. As seen in Figs. 2 and 3, there is a reasonable agreement between the predictions and the measurements. In particular, the various phases of the radiation mechanism are well simulated in both cases. Their origin is explained below.

In both cases (Figs. 2 and 3), the observation point was chosen close enough to the plate (around 10 cm) so that the acoustic radiation due to the initial deformation of the plate could be well separated from the radiation due to the waves propagating in the plate.

In the first case (Fig. 2 where $\theta=0°$ and $r=10$ cm), when the measurement point is near the normal to the plate at the impact point, the acoustic pressure time signal presents three different parts.

(i) First, a peak due to the direct contribution of the plate through its initial deformation at the impact point. This peak will be called the ''transient'' part of the time signal.

(ii) Second, a zone of low radiated pressure corresponding to the propagation of the bending waves in the plate away from the impact location. As the waves propagate in the plate away from the impact point, only the waves above the critical frequency will have a significant radiation level. As these waves above the critical frequency radiate at an angle between 0 and 90° in the direction of propagation away from the impact point, the radiated field will not reach the observation point located on the normal to the plate at the impact point.

(iii) Third, after these bending waves have reached the plate edges, have been reflected, and come back towards the impact point, the sound they radiate at an angle between 0 and 90° in the direction of propagation will now reach the measurement point located on the normal to the plate at the impact point. This part of the curve will be called the ''pseudo-permanent'' part of the time signal.

In the second case (Fig. 3, where $\theta=73°$ and $r=10.6$ cm), when the angle θ has a high value, the distinction between the three parts can no longer be made. Although the measurement point is still located at about the same distance of the impact location, it is now well off the normal to the impact point; hence, it will be reached sooner than previously by the radiation of the bending waves in the

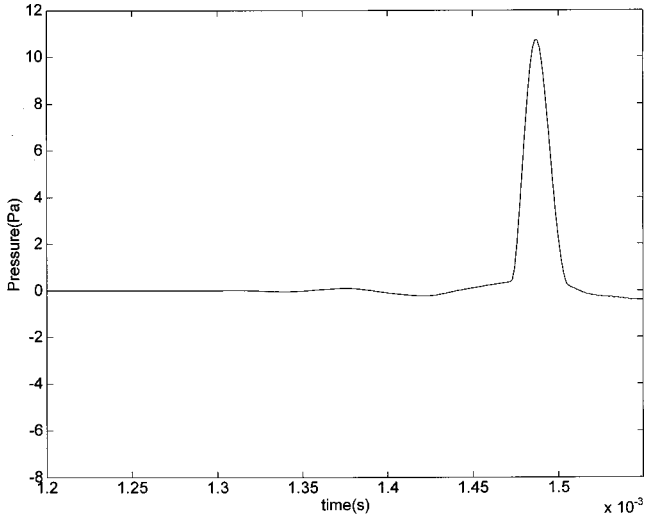


FIG. 4. Acoustic pressure for $\theta=30°$, $\varphi=0°$, $r=50$ cm, $V_0=1.4$ m/s, large aluminum plate (4.05×3.25×3 mm), steel sphere.

plate as their radiation will now reach the observation point on their first travel towards the plate edges (before their reflection at the plate edges). The radiation from the bending waves will reach the observation point before the radiation from the initial deformation of the plate and, consequently, the two radiation zones (transient and pseudo-permanent) will be superimposed.

A third case (Fig. 4 where $\theta=30°$ and $r=50$ cm) was simulated to show that even when the observation point is located at a greater distance from the plate (50 cm instead of 10 cm) and still fairly close to the normal to the plate at the impact point (30° angle with the normal), the peak due to the initial deformation can be well separated from the other radiation mode if the plate dimensions are made large enough (4.05×3.25 m instead of 0.405×0.325 m) to delay significantly the arrival at the observation location of the radiation due to the bending waves reflected at the edges (the figure shows only the peak due to the initial deformation as the second radiation mode has been delayed enough to lie outside the observation time window).

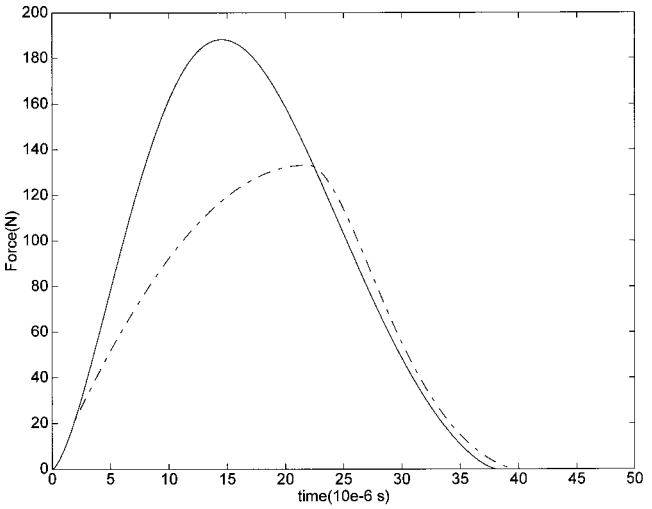


FIG. 5. Impact force, $V_0=1.4$ m/s, standard aluminum plate, steel sphere, elastic (solid) and elasto-plastic (dashed) cases.

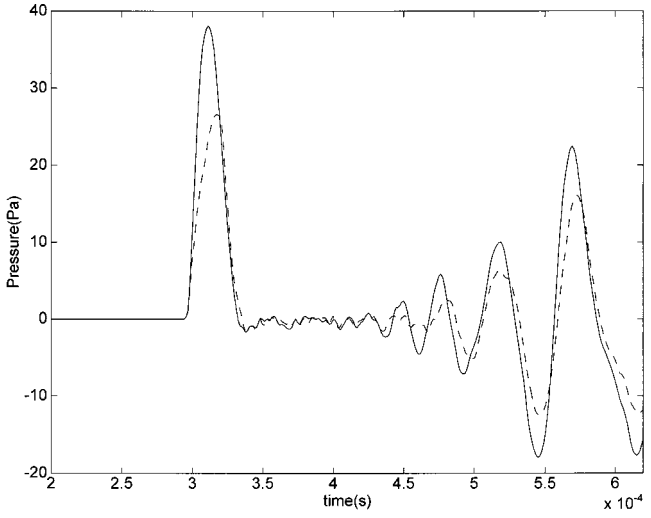


FIG. 6. Acoustic pressure, $V_0=1.4$ m/s, standard aluminum plate, steel sphere, elastic (solid) and elasto-plastic (dashed) cases, $\theta=\varphi=0°$, $r=10$ cm, $\xi=0.025$.

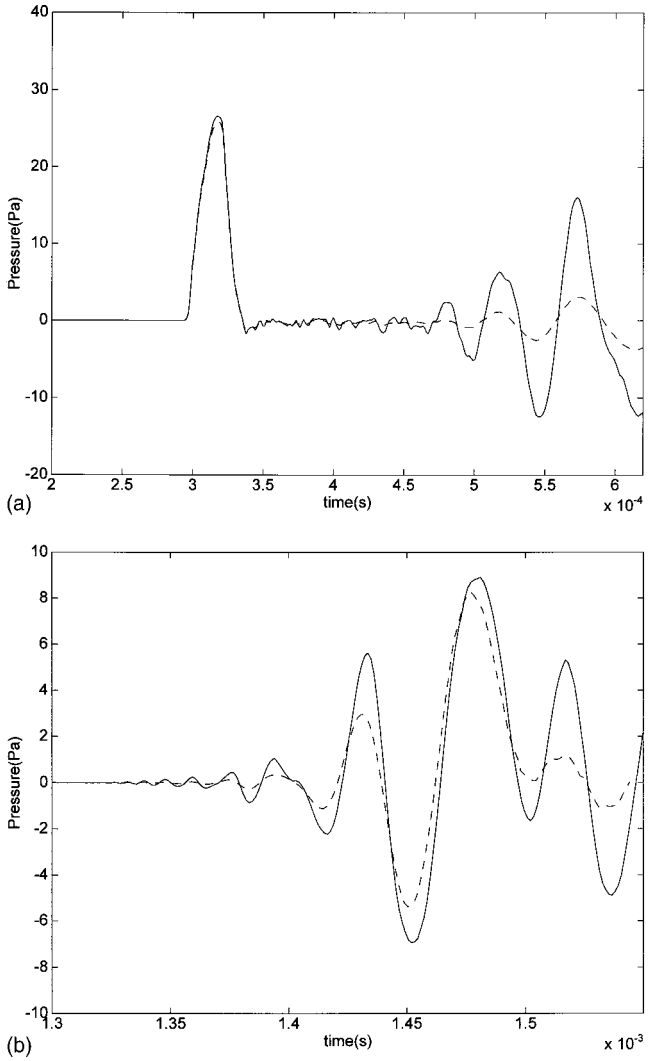


FIG. 7. Acoustic pressure, $V_0=1.4$ m/s, standard aluminum plate, influence of the damping ratio: $\xi=0.025$ (solid) and $\xi=0.1$ (dashed). (a) $\theta=\varphi=0°$, $r=10$ cm; (b) $\theta=30°$, $\varphi=0°$, $r=50$ cm.

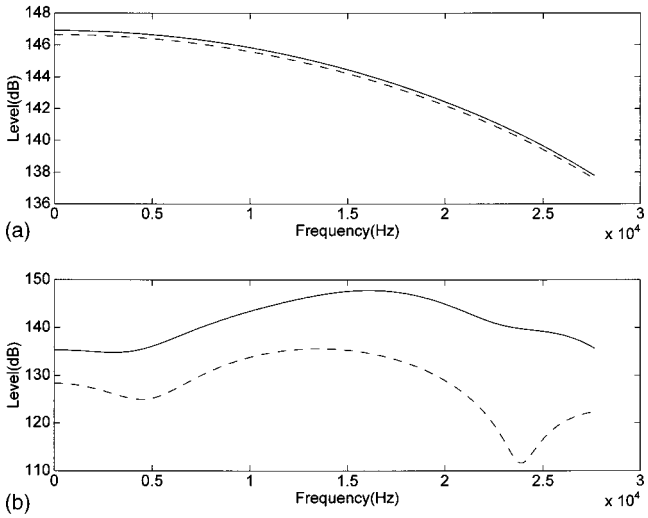


FIG. 8. Pressure spectrum, case of Fig. 6(a). (a) Transient part of the time-pressure curve, $\xi=0.025$ (solid) and $\xi=0.1$ (dashed); (b) Pseudo-permanent part of the time-pressure curve, $\xi=0.025$ (plain) and $\xi=0.1$ (dashed).

IV. PARAMETRIC STUDY

In this section, the numerical model is used to study the influence of plasticity, damping, materials, and impact velocity on sound radiation.

Taking into account plasticity in the collision causes the maximum value of the impact force to be lower and later in time, when compared to the case of a perfectly elastic (Hertzian) collision. The duration of contact is however about the same (Fig. 5). Figure 6 shows that taking into account the plasticity of the plate material induces a global reduction of the pressure level for both the transient and the pseudo-permanent zones.

Damping reduces significantly the acoustic pressure only in the pseudo-permanent zone [Fig. 7(a)]. Figure 7(b) demonstrates that the action of damping remains the same in the case of a microphone position where the transient and pseudo-permanent zones are superimposed, as the small decrease of the larger peak is due to the fact that this peak

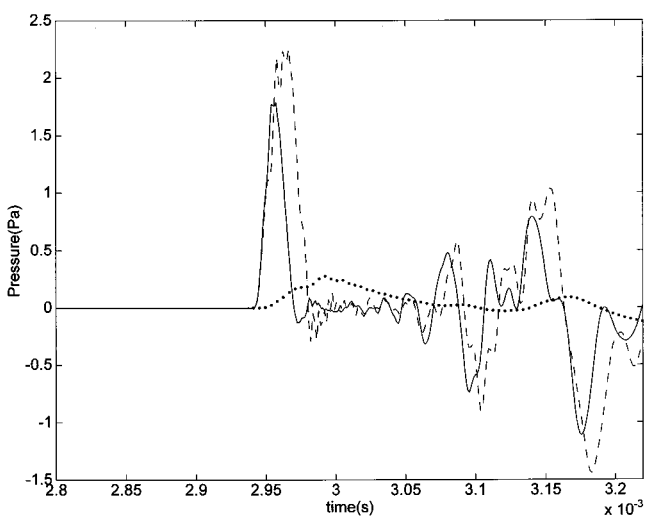


FIG. 9. Acoustic pressure, $V_0=1.0$ m/s, $\theta=\varphi=0°$, $r=1.0$ m, $\xi=0.025$, steel sphere, influence of plate material: steel (solid), aluminum (dashed), Plexiglas (dotted).

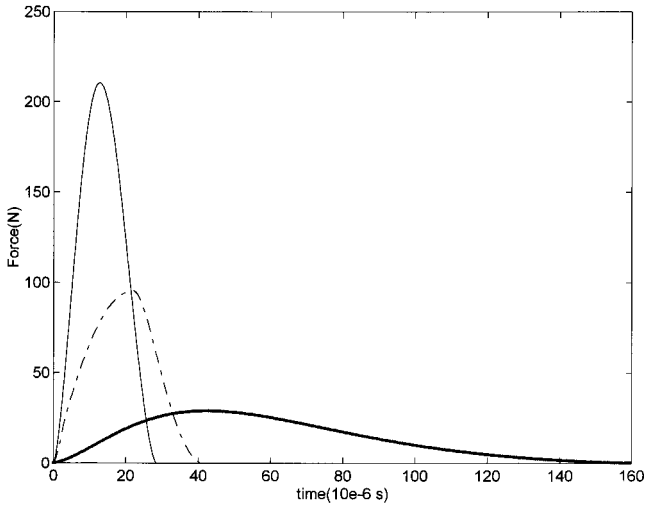


FIG. 10. Impact force, $V_0=1.0$ m/s, $\xi=0.025$, steel sphere, influence of plate material: steel (solid), aluminum (dashed), Plexiglas (dotted).

depends mainly on the contribution of the transient part of the signal. Separate Fourier transforms of the transient and pseudo-permanent zones of Fig. 7(a) are given, respectively, in Figs. 8(a) and (b). It confirms that damping affects only the pseudo-permanent zone.

Changing materials of impacting structures may be a way to reduce noise. In terms of an industrial case, only the plate material may be modified (that is, the impacted structure), considering that the sphere material (which could be an impacting tool) is a constraint that cannot be changed. Figure 9 presents the pressure radiated when a steel sphere collides with a steel plate, an aluminum plate, and a Plexiglas plate. Plasticity of the plate is taken into account. The measurement point is at $\theta=\varphi=0°$ and $r=1$ m, and the initial velocity of the sphere is 1 m/s. Considering the acoustic pressure in Fig. 9 and the corresponding impact forces in Fig. 10, it can be seen that the shape and magnitude of the time-force curves will directly influence the time-pressure curves, but that for two different materials, a lower force does not nec-

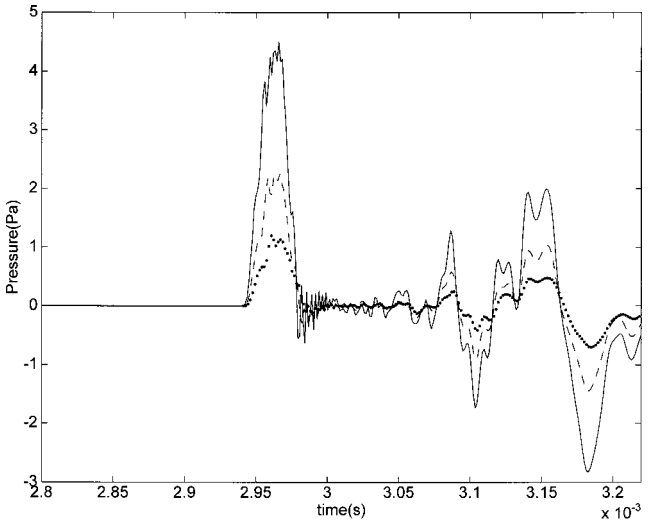


FIG. 11. Acoustic pressure, $\theta=\varphi=0°$, $r=1.0$ m, $\xi=0.025$, steel sphere, aluminum plate, influence of impact velocity: $V_0=0.5$ m/s (dotted), $V_0=1.0$ m/s (dashed), $V_0=2.0$ m/s (solid).

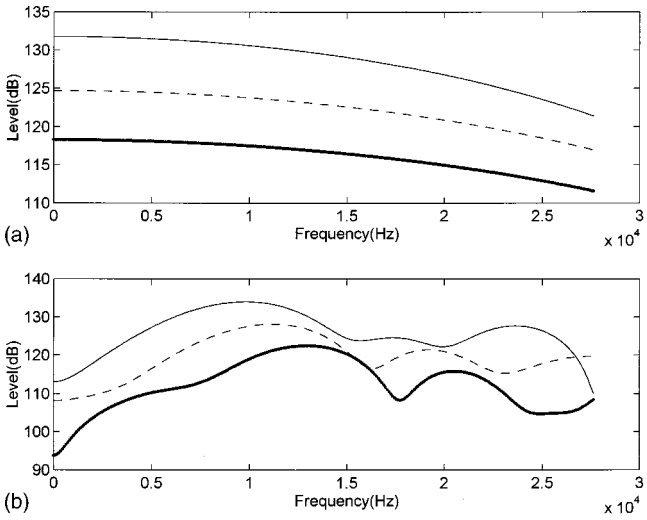


FIG. 12. Pressure spectrum, case of Fig. 10. (a) Transient part of the time-pressure curves, $V_0=0.5$ m/s (dotted), $V_0=1.0$ m/s (dashed), $V_0=2.0$ m/s (solid); (b) Pseudo-permanent part of the time-pressure curves, $V_0=0.5$ m/s (dotted), $V_0=1.0$ m/s (dashed), $V_0=2.0$ m/s (solid).

essarily imply less radiation; the aluminum plate radiates more than the steel plate, although the impact force is weaker. This can be explained by the fact that for the same thickness and lateral dimensions, the steel plate has a much more important mass than the aluminum plate, and Eq. (7) shows that the displacement of the plate is proportional to the force but also to the inverse of the mass.

The initial velocity of the sphere also has an influence on the pressure radiated (Fig. 11). The radiation spectrum at a point defined by $\theta=\varphi=0°$ and $r=1$ m [Figs. 12(a) and (b)] shows that each doubling of the initial velocity (from 0.5 to 1 m/s and from 1 to 2 m/s) globally increases the pressure level by 6 dB in both the transient and the pseudo-permanent zones, confirming the results of Ref. 10.

V. CONCLUSIONS

A comprehensive model for time-domain predictions of the acoustic radiation from a simply supported rectangular plate impacted by a sphere has been developed and validated experimentally.

The model was used to illustrate how the acoustic radiation is composed of two independent sources; the initial deformation of the plate and the propagation of bending waves in the plate with reflections at the edges of the plate.

Its use as a tool for the acoustical design of mechanical systems whose operation includes impacts was illustrated with a parametric study on plasticity, damping, plate materials, and impact velocity. Plasticity has the effect of reducing the sound level from both sources of radiation. Damping has the effect of reducing the sound level due to bending waves propagation only, the sound level due to initial plate deformation being practically unaffected. The effect of plate materials does not follow a simple trend, as a lower force resulting from a change in material does not necessarily produce less radiation. For the initial impact velocity, a 6-dB increase in sound level for each doubling of the impact velocity was confirmed.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of NSERC (Natural Sciences and Engineering Research Council of Canada) and FCAR (Fonds pour la formation de Chercheurs et l'Aide à la Recherche).

[1] W. Goldsmith, *Impact* (Edward Arnold, London, 1960).
[2] D. Tabor, *The Hardness of Metals* (Clarendon Press, Oxford, 1951).
[3] A. W. Crook, ''A study of some impacts between metal bodies by a piezo-electric method,'' Proc. R. Soc. London, A **212**, 377–390 (1952).
[4] R. M. Davies, ''The determination of static and dynamic yield stresses using a steel ball,'' Proc. R. Soc. London, Ser. A **197**, 416–432 (1949).
[5] S. Chattopadhyay, ''Permanent indentation effects on the impact response of elastic plates,'' J. Acoust. Soc. Am. **82**, 493–497 (1987).
[6] L. L. Koss and R. J. Alfredson, ''Transient sound radiated by spheres undergoing an elastic collision,'' J. Sound Vib. **27**(1), 59–75 (1973).
[7] L. L. Koss, ''Transient sound from colliding spheres—inelastic collisions,'' J. Sound Vib. **36**(4), 555–562 (1974).
[8] A. Akay and M. Latcha, ''Sound radiation from an impact-excited clamped circular plate in an infinite baffle,'' J. Acoust. Soc. Am. **74**, 640–648 (1983).
[9] M. Tokunaga, A. Akay, and M. Latcha, ''A theoretical analysis of transient sound radiation from a clamped circular plate,'' J. Appl. Mech. **51**, 41–46 (1984).
[10] D. Takahashi, ''Frequency analysis of sound radiation from an impact-excited plate,'' J. Acoust. Soc. Am. **91**, 2708–2713 (1992).
[11] F. Laville and D. Trentin, ''Impact noise modeling with application to noise control at the source in the case of a plastic granulator,'' Inter-Noise 92, pp. 607–609.

Active control of coupled structural/acoustic intensities in a fluid-loaded elastic plate

Matthias Fischer[a] and Sabih I. Hayek
Active Vibration Control Laboratory, Department Engineering Science and Mechanics, 212 Earth and Engineering Sciences Building, Penn State University, University Park, Pennsylvania 16802

(Received 26 May 1999; accepted for publication 10 July 2000)

In this paper, active control of structural intensity and radiated acoustic power in a coupled structure–fluid system is explored. The response of an infinite fluid-loaded plate and the resultant acoustic pressure field are calculated using a hybrid analytic/numerical approach. Two control strategies are examined: (1) the minimization of the total input power, and (2) the minimization of a weighted sum of structural intensity and the radiated acoustic power. Below coincidence, the vibration response of the plate is dominated by nondecaying waves, thus it can be reduced very efficiently by the active control strategies of minimizing the total input power or minimizing the structural intensity. Below coincidence, the acoustic radiation is caused by fast decaying structural waves thus, and active control strategies that aim to minimize the structural intensity have little impact on the radiated pressure field. Furthermore, the control strategy for minimizing the radiated acoustic power does not perform satisfactorily below coincidence since the controllers cannot be coupled to the fast decaying waves. Above coincidence, a slowly decaying structural wave dominates both the plate response and the acoustic radiation, and thus the structural intensity and the radiated acoustic power can be reduced significantly by each of the two control strategies. [S0001-4966(00)00211-3]

PACS numbers: 43.40.Vn, 43.40.Rj, 43.50.Ki [PJR]

I. INTRODUCTION

Structural vibrations can cause serious problems in many applications and active vibration control has become one of the methods to minimize vibrational response. One active control concept is to minimize the acceleration of the structure at some sensor point(s). A second concept is to minimize the power flow in a zone or in the entire structure. The control of the power flow is much costlier than simple acceleration control since power cannot only flow in the form of shear waves in the structure, but also in the form of bending and torsional waves. The cost function to be minimized for the power flow control strategy is the structural intensity on a defined structural path. Thus power flow control has advantages in applications where the propagation of the vibration in a structure is to be minimized.

The first attempt to measure the structural intensities in beams and plates was proposed by Noiseux.[1] The control of mechanical power flow in structures was examined widely in the past ten years. Fuller and Gonidou[2] and Pan and Hansen[3] developed active control strategies for minimization of mechanical power flow across an error sensor in beams. Pan and Hansen[4] compared the results obtained for mechanical power flow control and acceleration control. Schwenk *et al.*[5] developed an active control algorithm to experimentally control the total structural intensity in a finite beam with a complex impedance termination.

The important aspect of controlling the total input power supplied to an unloaded structure was investigated by Jenkins *et al.*,[6] Pavic,[7] and Bardou *et al.*[8]

To obtain the response and the power flow of structures with heavy fluid-loading, one has to solve a coupled system. By using Fourier transforms, one can obtain the solution in the wave number domain. Several authors have investigated numerical methods to calculate the inverse transform for this problem. Feit and Liu[9] obtained the response of a line-driven fluid-loaded elastic plate using a numerical contour integral approach. Seren and Hayek[10] compared different numerical methods to solve the Fourier integral for an insonified elastic plate with a line discontinuity. They proposed to subtract all the poles of the Fourier integrand to obtain a fast converging numerical solution. Cuschieri and Feit[11] used fast Fourier transform algorithms to evaluate the Fourier integral after subtracting the two real poles of the integrand.

Active control of structures submerged in a heavy acoustic medium has been discussed by few authors. Gu and Fuller,[12] Meirovitch,[13] and Lee and Park[14] all limited their examination to the minimization of the radiated noise. In this paper, both aspects, i.e., the control of the structural intensity and the radiated acoustic power, are considered.

Before using active control one has to define the design goals that should be achieved. If the designer of a control system is interested in keeping the operating energy costs of the system as low as possible, then minimization of total input power would be an efficient control strategy. In another design strategy, a specific area needs to be isolated from harmful vibrations and thus the structural intensity, i.e., the power flow to that zone, has to be minimized. In yet a third important design strategy the goal could be to minimize the radiated acoustic power.

In this paper, the efficiency of those control strategies with respect to their objective target is examined. Further-

[a]Currently at Institut A für Mechanik, Universität Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany.

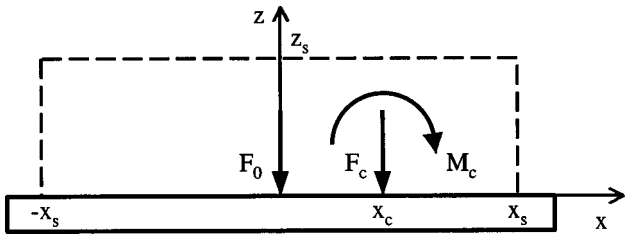


FIG. 1. Fluid-loaded plate.

more, their effect on secondary goals is studied. This leads to a compromise strategy, where structural power flow and radiated acoustic power can be controlled simultaneously.

II. PLATE RESPONSE AND RADIATED PRESSURE

An infinite elastic plate ($-\infty<\tilde{x}<\infty$) is loaded with a fluid on the upper side ($\tilde{z}>0$). The plate is excited by a time-harmonic line-force $\tilde{F}_0 e^{-i\omega t}$ located at the origin $\tilde{x}=0$. A time-harmonic line-force $\tilde{F}_c e^{-i\omega t}$ and a line-moment $\tilde{M}_c e^{-i\omega t}$ are applied as controllers, located at $\tilde{x}_c$ (see Fig. 1). The harmonic time dependency $e^{-i\omega t}$ is suppressed from this point on. The response of the system is calculated following the approach of Cuschieri and Feit.[11] Using the Bernoulli–Euler plate theory, the displacement of the plate $\tilde{w}(\tilde{x})$ is described by

$$D\left[\frac{d^4}{d\tilde{x}^4}-\alpha^4\right]\tilde{w}=-\tilde{p}(\tilde{x},\tilde{z}=0)+\tilde{F}_0\delta(\tilde{x})+\tilde{F}_c\delta(\tilde{x}-\tilde{x}_c)+\tilde{M}_c\delta_1(\tilde{x}-\tilde{x}_c), \tag{1}$$

where $D=Eh^3/12/(1-\nu^2)$ is the bending stiffness of the plate, α is the flexural wave number given by $\alpha^4=\rho_1 h\omega^2/D$, $\tilde{p}$ is the acoustic pressure, and δ and δ_1 are the Dirac delta functions of zero and first order, respectively. Equation (1) can be normalized:

$$\left[\frac{d^4}{dx^4}-1\right]w=-p(x,z=0)+\frac{F_0}{\alpha}\delta\left(\frac{x}{\alpha}\right)+\frac{F_c}{\alpha}\delta\left(\frac{x-x_c}{\alpha}\right)+\frac{M_c}{\alpha^2}\delta_1\left(\frac{x-x_c}{\alpha}\right), \tag{2}$$

where $x=\alpha\tilde{x}$, $z=\alpha\tilde{z}$, $w=\alpha\tilde{w}$, $p=\tilde{p}/(D\alpha^3)$, $F_0=\tilde{F}_0/(D\alpha^2)$, $F_c=\tilde{F}_c/(D\alpha^2)$, and $M_c=\tilde{M}_c/(D\alpha)$. The normalized differential equation for the acoustic pressure field $p(x,z)$ is given by

$$\left[\frac{\partial^2}{\partial z^2}+\frac{\partial^2}{\partial x^2}+\left(\frac{k_0}{\alpha}\right)^2\right]p=0, \tag{3}$$

where $k_0=\omega/c_0$ is the acoustic wave number. The pair of complex Fourier transforms on x is defined by

$$F(k)=\int_{-\infty}^{\infty}f(x)e^{-ikx}\,dx,\qquad f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}F(k)e^{ixk}\,dk. \tag{4}$$

Applying the transform on the coupled differential equations, Eqs. (2) and (3), and using the scaling law of the Dirac delta functions, the governing equations on the transformed displacement W and pressure P in the wave number domain are obtained as

$$(k^4-1)W=-P(k,z=0)+F_0+(F_c+ikM_c)e^{ikx_c}, \tag{5}$$

$$\frac{d^2P}{dz^2}-\left(k^2-\left(\frac{k_0}{\alpha}\right)^2\right)P=0. \tag{6}$$

Continuity of the normal velocities of the plate and the acoustic medium at the surface leads to

$$W(k)=\frac{1}{Z_s+Z_a}(F_0+(F_c+ikM_c)e^{ix_ck}). \tag{7}$$

The normalized structural and acoustic displacement impedances Z_s and Z_a, respectively, are defined as

$$Z_s=k^4-1=-i\omega\tilde{Z}_s,\quad Z_a=\frac{\epsilon}{\sqrt{\Omega}}\frac{1}{i\sqrt{\Omega-k^2}}=-i\omega\tilde{Z}_a, \tag{8}$$

where $\tilde{Z}_s$ and $\tilde{Z}_a$ are the structural and acoustic impedances, respectively. $\Omega=\omega/\omega_c$ and $\epsilon=\rho_0 c_0/(\rho_1 h\omega_c)$ are the frequency and fluid loading parameters, respectively. The coincidence frequency $\omega_c=c_0^2\sqrt{\rho_1 h/D}$ is defined when the flexural wave number α in the plate equals the acoustic wave number k_0 in the fluid. The coincidence frequency is an important transition point for fluid-loaded structures, since below coincidence, straight crested structural waves do not radiate acoustic power to the acoustic farfield. For frequencies above coincidence, straight crested waves radiate acoustic power to the fluid.

The radiated pressure is found in the wave number domain as

$$P(k,z)=\frac{Z_a}{Z_s+Z_a}(F_0+(F_c+ikM_c)e^{ix_ck})e^{iz\sqrt{\Omega-k^2}}. \tag{9}$$

The real part of the structural intensity vector represents the time-averaged energy flow in the plate. Shear forces, bending moments, and twisting moments usually contribute to the structural intensity. Since in this work only line-forces and line-moments act on the plate, the problem collapses to one dimension. The structural intensity in this case is defined by

$$SI(\tilde{x})=D\left(\frac{\partial^3\tilde{w}}{\partial\tilde{x}^3}\frac{\partial\tilde{w}^*}{\partial t}-\frac{\partial^2\tilde{w}}{\partial\tilde{x}^2}\frac{\partial^2\tilde{w}^*}{\partial\tilde{x}\,\partial t}\right), \tag{10}$$

where $()^*$ denotes the complex conjugate. The first term in Eq. (10) represents the structural intensity in the form of shear force times the plate velocity; the second term is the contribution of the bending moment times the angular velocity. In normalized variables and for the harmonic time dependency $e^{-i\omega t}$, Eq. (10) can be written as

$$SI(x)=i\rho_1 h(c_0\sqrt{\Omega})^3\left(\frac{\partial^3 w}{\partial x^3}w^*-\frac{\partial^2 w}{\partial x^2}\frac{\partial w^*}{\partial x}\right). \tag{11}$$

One notes that for the calculation of the structural intensity, the plate response and its spatial derivatives up to order 3 are needed. The nth derivative of the plate displacement can be obtained by multiplying Eq. (7) by $(ik)^n$ and then applying the inverse Fourier transformation. This implies that the slope of the plate caused by a unit line-force is equal to the displacement obtained for forcing by a line-moment.

The inverse Fourier transformation of Eq. (7) can be performed using a hybrid analytical/numerical approach. This method uses numerically efficient fast Fourier transform (FFT) algorithms. The real poles of Eq. (7) represent singularities on the path of integration in the FFT and thus have to be subtracted analytically. Their contribution can be calculated using the residue theorem. The remainder function is analytic on the real axis and thus gives a fast converging solution in the FFT.

The poles of Eq. (7) are computed as the zeroes of $g(k)$:

$$g(k)=Z_s+Z_a=k^4-1+\frac{\epsilon}{\sqrt{\Omega}}\frac{1}{i\sqrt{\Omega-k^2}}=0. \tag{12}$$

The roots of the Eq. (12) can be found by squaring it and using a standard polynomial root finder. This process creates extra roots which do not satisfy Eq. (12). To find the correct roots, one has to specify the sign of the square root term. Exponent $iz\sqrt{\Omega-k^2}$ in Eq. (9) must have a negative real part and a positive imaginary part to allow for propagating and decaying pressure waves.

The loci of the correct roots show an interesting behavior at the coincidence frequency $\Omega=1$. For frequencies below coincidence, one finds two real roots $(\pm k_1)$ and two pairs of complex conjugate roots. For all of them, $|k|>\Omega$, so that they all represent subsonic waves in the plate, i.e., they propagate slower than the velocity of sound in the fluid. The real roots represent straight crested structural waves and hence do not radiate sound to the acoustic farfield. The two pairs of complex conjugate roots represent heavily damped structural waves and cause radiation of acoustic power to the acoustic medium. For forcing frequencies above coincidence, the real roots stay subsonic; the complex roots now represent supersonic waves in the plate. Two of the complex roots move close to the real axis. This pair now represents a slowly decaying wave in the plate.

By subtracting the contribution of the real roots $\pm k_1$, the Fourier integrals can be decomposed as follows:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{(ik)^n}{g(k)}e^{ixk}\,dk=\frac{1}{2\pi}\left(\int_{-\infty}^{\infty}\frac{2k_1}{g'(k_1)}\frac{(ik_1)^n}{k^2-k_1^2}\left(\frac{k}{k_1}\right)^{\eta}\times e^{ixk}\,dk+\int_{-\infty}^{\infty}M_n(k)e^{ixk}\,dk\right), \tag{13}$$

where $\eta=0$ for n even and $\eta=1$ for n odd. $g'(k)$ is found as

$$g'(k)=4k^3+\frac{\epsilon}{\sqrt{\Omega}}\frac{k}{i(\Omega-k^2)^{3/2}}. \tag{14}$$

The first integral in Eq. (13) can be solved analytically using residue theorem:

$$\int_{-\infty}^{\infty}\frac{2k_1}{g'(k_1)}\frac{(ik_1)^n}{k^2-k_1^2}\left(\frac{k}{k_1}\right)^{\eta}e^{ixk}\,dk=2\pi i\frac{(ik_1)^n}{g'(k_1)}e^{ik_1x}. \tag{15}$$

The remainder function $M_n(k)$ is given by

$$M_n(k)=\frac{(ik)^n}{g(k)}-\frac{2k_1}{g'(k_1)}\frac{(ik_1)^n}{k^2-k_1^2}\left(\frac{k}{k_1}\right)^{\eta}. \tag{16}$$

This is an analytic function on the real axis and thus gives a fast converging solution when used in an FFT algorithm. However, if $M_n(k)$ is evaluated numerically at $k=k_1$, the numerical subtraction will cause problems since the computer cannot perform limiting processes. Thus the limit $k\to k_1$ has to be found analytically using l'Hospital's rule, i.e.,

$$M_n(k_1)=\frac{(ik_1)^n}{4k_1g'(k_1)}\left(2-4(n-\eta)-\frac{2k_1}{g'(k_1)}g''(k_1)\right), \tag{17}$$

where $g''(k)$ is found as

$$g''(k)=12k^2+\frac{\epsilon}{i\sqrt{\Omega}}((\Omega-k^2)^{-3/2}+3k^2(\Omega-k^2)^{-5/2}). \tag{18}$$

The combined result for the normalized plate response is

$$\begin{aligned}\frac{\partial^n w}{\partial x^n}=&iF_0\frac{(ik_1)^n}{g'(k_1)}e^{ik_1x}+\frac{F_0}{2\pi}\int_{-\infty}^{\infty}M_n(k)e^{ikx}\,dk\\&+iF_c\frac{(ik_1)^n}{g'(k_1)}e^{ik_1(x-x_c)}+\frac{F_c}{2\pi}\int_{-\infty}^{\infty}M_n(k)e^{ik(x-x_c)}\,dk\\&+iM_c\frac{(ik_1)^{n+1}}{g'(k_1)}e^{ik_1(x-x_c)}\\&+\frac{M_c}{2\pi}\int_{-\infty}^{\infty}M_{n+1}(k)e^{ik(x-x_c)}\,dk.\end{aligned} \tag{19}$$

For forcing frequencies above coincidence, the two complex roots that are now close to the real axis have to be subtracted in addition to the real roots. These subtractions are performed in a similar manner to the subtraction of the real poles above in order to improve and accelerate the convergence of the FFT integral.

The numerical calculations in this paper were performed for a steel plate with water loading, i.e., a fluid-loading parameter of $\epsilon=0.132$. The results for forcing frequencies below and above coincidence are presented. Figure 2 shows the normalized displacement of the uncontrolled plate, i.e., forced by a unit line-force at $x=0$. Below coincidence, a frequency parameter of $\Omega=0.2$ was chosen as representative example. As expected the displacement has a maximum and the slope is zero at the position of the driving force. The displacement of the plate consists of two separate solutions as discussed earlier, the contribution of the real roots of Eq. (12) and the contribution of the FFT of the remainder function. Below coincidence, the real roots cause a nondecaying

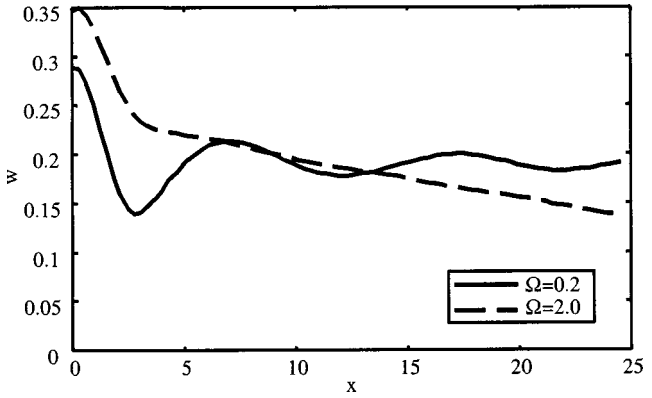


FIG. 2. Uncontrolled displacement field for $\Omega=0.2$ and $\Omega=2.0$.

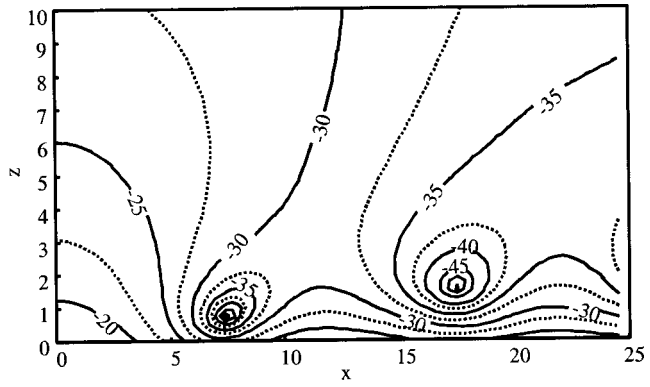


FIG. 4. Uncontrolled pressure field for $\Omega=0.2$.

structural wave which dominates the total plate response. The FFT contributes a fast decaying solution. The inverse FFT is performed with a sample space of 2×2048 points and with a wave number resolution of $\Delta k=0.01$.

For forcing frequencies above coincidence, a frequency parameter of $\Omega=2.0$ is chosen as a representative example. For frequencies above twice the coincidence frequency, one should be aware that the Bernoulli–Euler plate theory is no longer a good model, since it describes an infinitely increasing flexural phase velocity. A better approximation would be Mindlin's plate theory. The normalized displacement field for the uncontrolled plate at $\Omega=2.0$ is shown in Fig. 2. The nondecaying contribution of the still subsonic real roots is negligible above coincidence. Slowly decaying structural waves caused by the poles close to the real axis dominate the displacement field. The contribution of the inverse FFT of the remainder function which includes the complex roots with large imaginary part decays very fast. For observation points located more than one wavelength away from the source, this part of the solution is negligible.

Figure 3 shows the active structural intensity normalized with respect to the input power caused by a unit line-force at the origin. It should be noted that power flows in the positive x-direction for $x>0$ and in the negative x-direction for $x<0$. The jump in the structural intensity plot at $x=0$ thus corresponds to the input power injected into the plate.

The structural intensity, which is obtained from the displacement and its spatial derivatives, shows a behavior that corresponds to the displacement fields below and above coincidence. Below coincidence, there is a constant amount of energy flowing in the plate away from the source caused by the subsonic nondecaying waves represented by the real roots of Eq. (12). The decaying wave pattern of the combined waves represented by the two complex roots, causes interactions between the plate and the fluid. Energy is radiated to the acoustic field as well as put back into the plate. Above coincidence, the power flow in the plate decays slowly to a very small value that is determined by the contribution of the subsonic real roots.

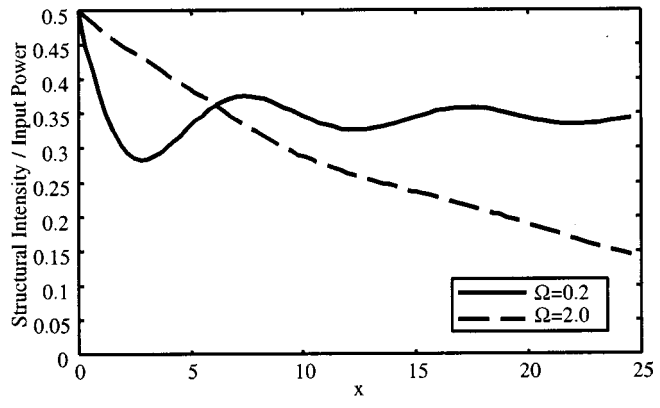


FIG. 3. Uncontrolled structural intensity for $\Omega=0.2$ and $\Omega=2.0$.

The inverse Fourier transform of Eq. (9) for the pressure field can be found using the same procedure as before for the plate response. In addition, the removable singularities at $k=\pm\sqrt{\Omega}$ have to be removed by subtracting the limit as $k^2\rightarrow\Omega$. The combined result for the pressure field caused by the driving force F_0 is obtained as

$$
\begin{aligned}
p(x,z)=&\frac{iF_0z\sqrt{\Omega}}{2r}H_1^{(1)}(r\sqrt{\Omega})\\
&-iF_0\frac{k_1^4-1}{g'(k_1)}e^{iz\sqrt{\Omega-k_1^2}}e^{ik_1x}\\
&-\frac{F_0}{2\pi}\int_{-\infty}^{\infty}\left(\frac{k^4-1}{g(k)}-\frac{2k_1}{g'(k_1)}\frac{k_1^4-1}{k^2-k_1^2}\right)\\
&\times e^{iz\sqrt{\Omega-k^2}}e^{ikx}\,dk,
\end{aligned}
\tag{20}
$$

where $r=\sqrt{x^2+z^2}$ is the normalized distance from the origin. For the control force and moment similar expressions can be derived. The last term in Eq. (20) can be evaluated using FFT algorithms again. But there is a major difference in the computational cost when compared to the calculation of the plate response. In that case each derivative could be evaluated by one transformation. On the other hand, to obtain the acoustic pressure, a transform has to be computed for each discrete value of z. Figure 4 shows the normalized pressure field in dB for forcing by a unit line-force at a frequency parameter $\Omega=0.2$. The maximum pressure is found at the driving point and there is a radiation peak in z-direction. One notices spots with very low pressure close to the plate surface. They are caused by the superposition of the nearfield

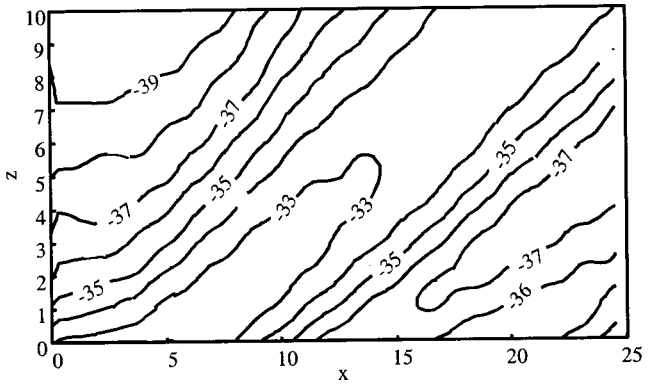


FIG. 5. Uncontrolled pressure field for $\Omega=2.0$.

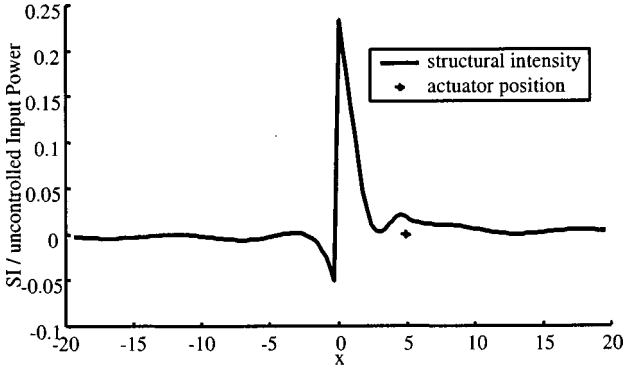


FIG. 6. Structural intensity for minimized input power, $\Omega=0.2$.

radiation of the nondecaying structural wave and the farfield radiation due to the complex roots. The normalized pressure field for a frequency parameter of $\Omega=2.0$ is shown in Fig. 5. The radiation peaks at an angle of 45° to the z-axis which corresponds to the coincidence angle for $\Omega=2.0$.

The real part of the acoustic intensity vector represents the time-averaged power flow through a unit area. For the harmonic time-dependency $e^{-i\omega t}$, it is given by

$$\mathbf{A}I=\frac{i}{\rho_0\omega}\tilde{p}\nabla\tilde{p}^*, \tag{21}$$

or in normalized variables:

$$\mathbf{A}I=\frac{i}{\rho_0}\sqrt{\frac{\rho_1 h}{D}}(mc_0^2\Omega)^2 p\nabla p^*. \tag{22}$$

The radiated acoustic power is obtained by integrating the normal component of the acoustic intensity vector along any closed path above the surface. For calculations in this paper, the boundary of a rectangular box measuring $2x_s$ by z_s (including all sources) was chosen as the path of integration, since the normal components then can be found easily by taking small differences of the pressure field. The radiated acoustic power is independent of the value of z_s where the contour is closed since there are no power sources or sinks away from the plate surface. The computations show no influence of the value of z_s when z_s is chosen larger than one flexural wavelength. For x_s larger than one flexural wavelength, the difference between the total power flow out of the box and the input power by the driving force is less than one percent of the input power.

III. ACTIVE CONTROL

In this paper, several control strategies are discussed. First, the results for forcing frequencies below coincidence are presented; the discussion of forcing frequencies above coincidence follows later on.

The first cost function is the minimization of the total input power supplied to the plate by both the primary force and the controllers:

$$C=F_0\frac{\partial w^*}{\partial t}(0)+F_c\frac{\partial w^*}{\partial t}(x_c)+M_c\frac{\partial^2 w^*}{\partial x\,\partial t}(x_c). \tag{23}$$

The plate response is calculated as discussed in Sec. II. The cost function is quadratic in the four unknown variables, the real and the imaginary part of the control force and the control moment. Thus the optimization can be performed analytically by setting the derivatives of the cost function to zero and solving the 4×4 system of linear algebraic equations for the complex control force and moment.

For a frequency parameter of $\Omega=0.2$ and a controller location of $x_c=5$, minimizing the total input power results in a reduction of the input power by a factor of 3. Figure 6 shows the normalized structural intensity for minimized input power. The plot is continuous at the controller location and thus there is no power injected into the plate by the controllers for this control strategy. The power flow in the plate outside the region between the primary force and controllers is practically zero; at the reference points $x=\pm10$ a reduction of 19.8 dB in structural intensity is achieved. Minimizing the total input power reduces the displacement of the plate, too. In Fig. 7 the absolute values of the displacement before and after control are shown. One notices a larger displacement of the controlled plate in the primary force/controller region, but a significant reduction of the displacement field by an average of 10 dB outside this region.

The influence of the location of the controllers x_c is significant. The best place to put the controllers is at the driving line. The control force is then the negative of the primary force, the control moment vanishes, and the total input power is zero. However, it is not always possible to place a controller exactly on the driving line, thus the total input power was calculated for values $0<x_c<12$. In Fig. 8

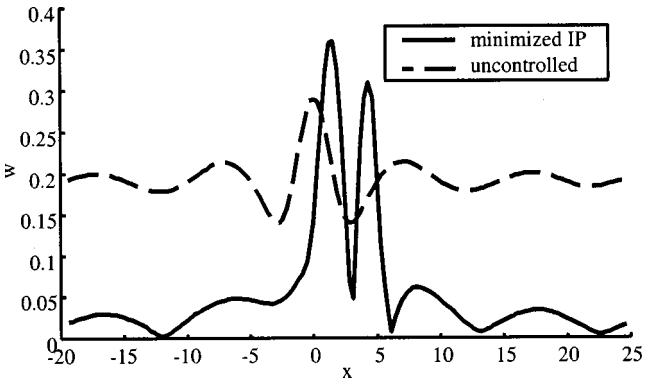


FIG. 7. Displacement for minimized input power, $\Omega=0.2$.

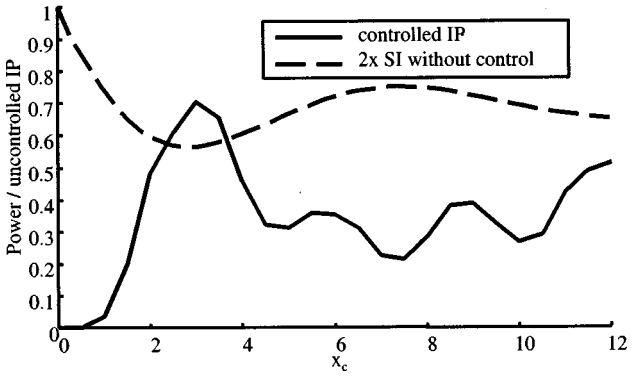


FIG. 8. Controlled input power vs x_c, input power minimized, $\Omega=0.2$.

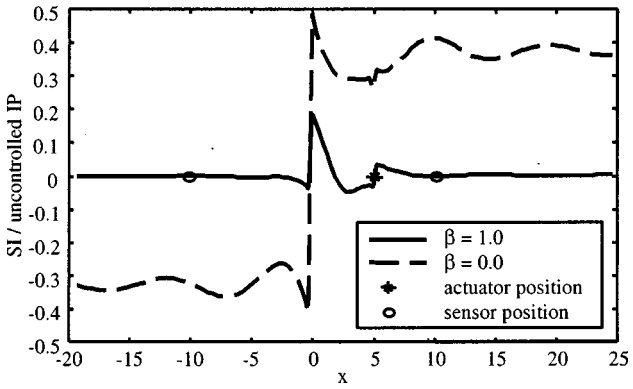


FIG. 9. Structural intensity for $\beta=1$ and $\beta=0$, $\Omega=0.2$.

these data are shown in comparison with twice the uncontrolled structural intensity at the controller location. The factor 2 was introduced to indicate the total outgoing power flow in the plate away from the source/controllers. The achieved reduction of input power varies widely with the controller location. For controllers at $x=0$, zero input power is achieved. The maximum input power for the controlled system is found as 0.75 of the uncontrolled input power. Two effects on the achieved reduction of input power can be noticed. One finds local minima and maxima for x_c at distances of approximately one-half of a flexural wavelength. This can be explained by the spatial phase difference between the primary force and the controllers. The second effect is that, in general, controller locations with a high uncontrolled power flow allow a more effective control.

Another control strategy is based on the minimization of structural intensity and radiated power in the system. One can define the cost function as a weighted sum of the structural intensity at two sensor points and the radiated power up to those points:

$$C=\beta SI+(1-\beta)RP, \tag{24}$$

where $SI=SI(x_s)-SI(-x_s)$, RP is the radiated acoustic power in the segment $-x_s$ to $+x_s$, and β is a weighting factor. To obtain the radiated power one has to add up the acoustic intensities in terms of the four unknown variables (real and imaginary parts of the control force and moment) along the chosen contour. That is a numerically tough job. A faster way to get approximately the same result is to calculate the radiated power by subtracting the structural intensity at the two sensor positions from the total input power (IP) due to the primary force and the controllers. The numerical calculations show that this method results in an error below 1%. The cost function becomes

$$C=(2\beta-1)SI+(1-\beta)IP. \tag{25}$$

Figure 9 shows the structural intensity plot for $\beta=1$ and $\beta=0$, $x_c=5$, and $\Omega=0.2$. A weighting factor of $\beta=1$ corresponds to a minimization of the total outgoing mechanical power in the plate at the two sensor points. The power flow in the plate can be reduced to a very small value outside the primary force/controller region by this strategy. At the sensor positions a structural intensity reduction of 26.1 dB is achieved. The results are very similar to those obtained by minimizing the total input power.

It should be mentioned at this point that the minimization of the structural intensity at one sensor point on the plate results in a complete suppression of power flow in this direction, but an increase of power flow in the opposite direction by a factor of 6.

The active control of the system with a weighting factor of $\beta=1$ has only little impact on the radiated acoustic field. The controlled normalized pressure field is shown in Fig. 10. The maximum pressure increases by 3.5 dB compared to the uncontrolled plate, but the radiated acoustic power does not change and amounts to 0.31× the uncontrolled input power for both cases.

For a weighting factor of $\beta=0$, one tries to minimize the radiated power between the sensors at $\pm x_s$. As shown in Fig. 9, this strategy has little impact on the plate response. Figure 11 shows the resultant pressure field. The maximum pressure increases by 0.5 dB when the $\beta=0$ strategy is applied. However, the radiated acoustic power can be reduced to 0.24× the uncontrolled input power.

For a compromise strategy of controlling a weighted sum of structural intensity and radiated power, one can choose a weighting factor $0<\beta<1$. Starting with $\beta=1$ and decreasing β, the results change only very slightly. At $\beta=0.5$, one weighs the structural and acoustic response equally which corresponds to the minimization of the total input power, with results approximately the same as for $\beta=1$. Decreasing β further, the structural intensity at the sen-

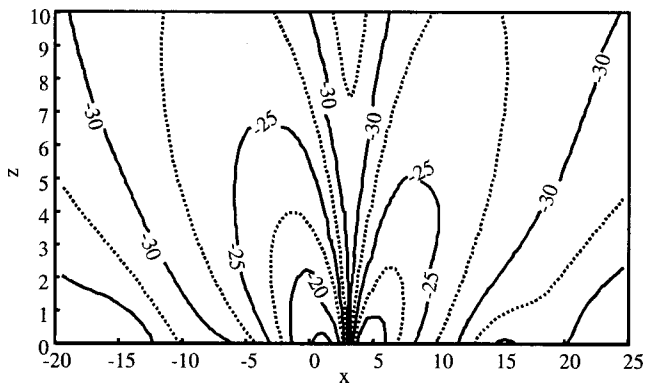


FIG. 10. Pressure field for $\beta=1$, $\Omega=0.2$.

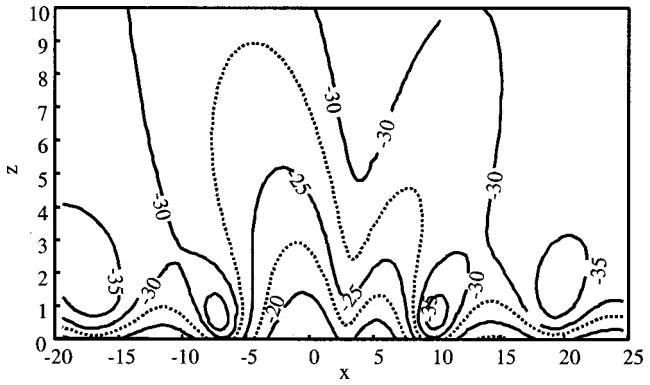


FIG. 11. Pressure field for $\beta=0$, $\Omega=0.2$.

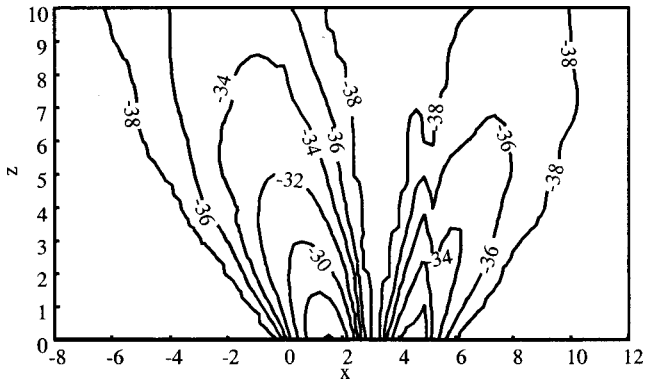


FIG. 13. Pressure field for $\beta=1$, $\Omega=2.0$.

sor points and the total input power start to increase rapidly and the radiated power decreases slowly.

In Fig. 12, the structural intensity plots for minimizing the total input power and the $\beta=1$ and $\beta=0$ strategies above coincidence are shown. The controller location is $x_c=5$ and the frequency parameter $\Omega=2.0$. For forcing frequencies above coincidence, the results for controlling the total input power and the structural intensity in the plate do not change significantly compared to control below coincidence. The power flow in the plate outside the primary force/controller region is very small; it is reduced at the positions $x=\pm 10$ by 17.1 dB and 38.8 dB for minimizing total input power and the $\beta=1$ strategy, respectively. The controlled total input and radiated power drop to 0.25 of the uncontrolled input power, since above coincidence all input power gets radiated to the acoustic medium with or without control. The controlled pressure field is shown in Fig. 13. The maximum pressure increases by 5.35 dB and 6.25 dB for minimizing the total input power and the structural intensity at the sensor points, respectively.

Again, a significant dependence on the controller location x_c is found. Above coincidence, the structural intensity in the uncontrolled plate is decreasing continuously, and thus the active control strategies become less and less efficient for increasing distance between source and controllers.

For the control strategy of minimizing the radiated acoustic power between the sensor points, the results look different for forcing frequencies above coincidence. The $\beta=0$ strategy is found to reduce input power and plate response above coincidence, too. The structural intensity plot (Fig. 12) shows a decaying power flow in the plate that is reduced by 10.7 dB at the sensor points. The total input power is reduced to $0.25\times$ the uncontrolled input power. The controllers in this case absorb power from the system. The controlled pressure field is similar to the $\beta=1$ case, the maximum pressure increases by 4.5 dB from the uncontrolled value, and the radiated acoustic power between the sensor points is $0.21\times$ the uncontrolled input power. The acoustic power that is radiated by the entire plate drops by a factor of 4 for control with $\beta=0$ when compared to the uncontrolled plate.

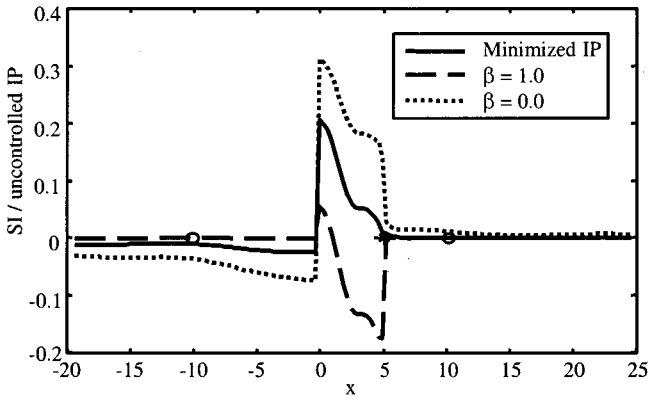


FIG. 12. Structural intensity for minimized input power, $\beta=1$ and $\beta=0$, $\Omega=2.0$.

IV. CONCLUSION

The response of an infinite elastic plate submerged in a heavy acoustic medium and the resultant acoustic pressure field were computed using a hybrid analytical/numerical approach. From the plate response, the mechanical power flow in the plate and the radiated acoustic power were calculated and used as cost functions in the active control strategies discussed in this paper.

The behavior of the plate is different for forcing frequencies below or above coincidence. For frequencies below coincidence, there is a nondecaying structural wave traveling in the plate. The strong interaction of the structure and the fluid causes power to be continuously radiated to the acoustic medium and vice versa. For forcing frequencies above coincidence, the contribution of the nondecaying wave in the plate is negligible and thus all input power into the plate radiates to the acoustic medium.

The control strategies that try to minimize total input power or structural intensity in the plate worked efficiently for all forcing frequencies. For forcing frequencies below coincidence, the radiated acoustic power could only be reduced slightly with the chosen controller configuration. For frequencies above coincidence, the acoustic radiated power is reduced significantly for minimizing the radiated power, the total input power into the system, or the structural intensity in the plate.

The results for forcing frequencies below coincidence show an interesting behavior. Without control, most of the input power flows in the plate in the form of nondecaying flexural waves. Thus one can get efficient control of the plate

response even with controllers placed at large distances from the source. For frequencies above coincidence the structural intensity in the plate decays monotonically and power is radiated to the acoustic field. Thus it is important to place controllers and sensors as close as possible to the source. For distances above five flexural wavelengths, the source and the controllers are practically decoupled and control is not possible.

It should be pointed out that for frequencies below coincidence, the control strategies of minimizing total input power or minimizing structural intensity only reduce the structural response of the plate. They have small effect on the radiated acoustic power. On the other hand, minimizing the radiated acoustic power does not reduce the vibration response of the plate, but only decreases the radiated power. It is not possible to get a significant improvement in both the structural intensity and the acoustic radiated power. This is because, below coincidence, the vibration of the plate and the acoustic radiation are dominated by two different structural waves. The waves represented by the real and complex roots cannot be controlled simultaneously with the controller configuration that is examined in this paper. Moreover, the wave responsible for the acoustic radiation decays very fast in the plate, and thus an effective reduction of the radiated acoustic power is not possible with controllers located at some distance of the forcing. Below coincidence, there is little energy in the supersonic region of the wave number spectrum for control of the acoustic radiated power to be effective. If the plate is finite, there would be additional scattering in the wave number spectrum into the supersonic region to make control more effective. For forcing frequencies above coincidence, both the structural and acoustic response are dominated by a slowly decaying wave. Control of the structure also decreases the radiated acoustic power and vice versa.

[1] D. U. Noiseux, ''Measurement of power flow in uniform beams and plates,'' J. Acoust. Soc. Am. **47**, 238–247 (1970).
[2] C. R. Fuller and L. O. Gonidou, ''Active vibration control of flexural power flow in beams,'' J. Acoust. Soc. Am. **84**, S47 (1988).
[3] J. Pan and C. H. Hansen, ''Active control of total power flow along a beam. I: Physical system analysis,'' J. Acoust. Soc. Am. **89**, 200–209 (1990).
[4] X. Pan and C. H. Hansen, ''Effect of error sensor location and type on the active control of beam vibration,'' J. Sound Vib. **165**, 497–510 (1993).
[5] A. E. Schwenk, S. D. Sommerfeldt, and S. I. Hayek, ''Adaptive control of structural intensity associated with bending waves in a beam,'' J. Acoust. Soc. Am. **96**, 2826–2835 (1994).
[6] M. D. Jenkins, P. A. Nelson, R. J. Pinnington, and S. J. Elliott, ''Active isolation of periodic machinery vibrations,'' J. Sound Vib. **166**, 117–140 (1993).
[7] G. Pavic, ''Input power suppression of vibrating infinite beams and plates by secondary sources,'' Proc. Inst. Acoustics **15**, 395–402 (1993).
[8] O. Bardou, P. Gardonio, S. J. Elliott, and R. J. Pinnington, ''Active power minimization and power absorption in a plate with force and moment excitation,'' J. Sound Vib. **208**, 111–151 (1997).
[9] D. Feit and Y. N. Liu, ''The near field response of a line-driven fluid-loaded plate,'' J. Acoust. Soc. Am. **78**, 763–766 (1985).
[10] C. Seren and S. I. Hayek, ''Acoustic radiation from an insonified elastic plate with a line discontinuity,'' J. Acoust. Soc. Am. **86**, 195–209 (1989).
[11] J. M. Cuschieri and D. Feit, ''A hybrid numerical and analytical solution for the Green's function of a fluid-loaded elastic plate,'' J. Acoust. Soc. Am. **95**, 1998–2005 (1994).
[12] Y. Gu and C. R. Fuller, ''Active control of sound radiation from fluid-loaded rectangular plate,'' J. Acoust. Soc. Am. **93**, 337–345 (1993).
[13] L. Meirovitch, ''Theory for the optimal control of the far-field acoustic pressure radiating from submerged structures,'' J. Acoust. Soc. Am. **93**, 356–362 (1993).
[14] H.-K. Lee and Y.-S. Park, ''A near-field approach to active control of sound radiation from a fluid-loaded rectangular plate,'' J. Sound Vib. **196**, 579–593 (1996).

Reducing seat dip attenuation

W. J. Davies and T. J. Cox
School of Acoustics and Electronic Engineering, Salford University, Salford M5 4WT, United Kingdom

(Received 4 January 2000; revised 1 June 2000; accepted 18 August 2000)

Strategies for reducing seat dip attenuation in concert halls are considered. It is shown that the dip is established 4 ms after the direct sound from the stage arrives at the listener. Sound scattered from the seats and floor is the main cause of the dip. By controlling these very early reflections the attenuation can be reduced to below its subjective threshold. With this in mind, changes to the shape and impedance of the seats and floor are trialed using a boundary element model and a physical scale model. It is found that the seat dip effect can be rendered inaudible by introducing a 1-m pit under the seats. Smaller improvements are produced by changing the impedance of the seat squab.
 [S0001-4966(00)04511-2]

PACS numbers: 43.55.Br, 43.55.Fw, 43.55.Ka, 43.20.Fn [JDQ]

I. INTRODUCTION

This article is concerned with seat dip attenuation, the anomalous low-frequency attenuation suffered by sound traveling at grazing incidence over rows of seats. The effect was first reported by two teams in 1964,[1,2] during an investigation of the poor acoustics of the New York Philharmonic Hall. The reports communicated the alarming fact that the attenuation could be as severe as 20 dB around 150 Hz. A larger attenuation will be observed if the angle of incidence approaches closer to grazing, or the sound travels over more rows of seats; a smaller attenuation will be observed if the receiver height is increased.

Since these first papers, there have been further explorations of the effect and yet complete agreement on the cause of seat dip attenuation does not yet exist. The reports of Sessler and West and of Schultz and Watters concluded that the effect seemed mainly due to a vertical resonance in the gaps between the rows of seats. This frequency domain explanation has been followed by Bradley[3] who argued for both vertical and horizontal resonances. However, frequency domain models do not explain every aspect of the seat dip effect. For example, the attenuation changes over time in the very early sound field.[4] Ishida *et al.*[5] were the first to explain seat dip effect in the time domain: many small reflections from the seats and floor produce a complicated impulse response immediately after the arrival of the direct sound from the stage. The seat dip attenuation is simply what results when this impulse response is Fourier transformed. More recent measurements have supported this idea.[6]

The frequency spectrum of the attenuation can be predicted accurately for a given configuration of source, seats, and receiver by a boundary element model (BEM).[7] A BEM is computationally expensive, however, and several other prediction schemes with varying levels of performance have been reported.[8–10] As desktop computing power increases the BEM is now probably the preferred prediction scheme.

Until recently, all investigations of the seat dip effect had concentrated on measuring, predicting, or explaining the dip. There were conjectures on the subjective significance of such a large attenuation, but it was not known whether the measured and predicted dips could actually be perceived by listeners. The subjective threshold of the effect has now been reported[11] to be a change of -3.8 ± 0.2 dB in the 200-Hz octave band early energy, from 0 to 80 ms. Thus the attenuation is likely to be audible, though not a disastrous demerit, in the majority of auditoria.

Three schemes have so far been suggested for reducing the attenuation. Ando *et al.*[8] used a numerical model to show that introducing resonant absorbers into the floor between rows of seats reduced the attenuation. In a later practical trial[4] in a real concert hall, it was found that resonant floor absorbers did reduce the attenuation but the benefit was insufficient to render it inaudible for most seats. The second remedial method suggested in the literature is to introduce strong early reflections which have not grazed seating. Bradley[3] reported that introducing a strong overhead reflection from ceiling reflectors reduced the seat dip attenuation in a hall. However, this method may introduce tonal colorations and may also reduce perceived spaciousness (as the proportion of lateral energy in the early field will be reduced). Finally, the attenuation can be reduced and its frequency shifted slightly by moving the vertical angle of incidence of the direct sound away from grazing.[4] However, the reduction is not great, so that a very steep seating rake will be needed to produce a significant decrease in the attenuation. At present, therefore, there is no fully effective method of reducing seat dip attenuation available to auditorium designers. This article reports on a series of experiments to reduce seat dip attenuation by altering the shape and impedance of the seats and floor, informed by a time domain view of the effect.

II. METHOD

A combination of scale modeling and BEM predictions were used to explore propagation over seating. The BEM predictions were chosen so that many different seat and floor geometries could be explored quite quickly. Scale model seats were used to explore the seat dip effect in the time domain, and also for confirmation of the BEM results. It should be noted that only the seats and the floor local to them were modeled. The results are thus for attenuation of the direct sound only and not for the total early sound field in an

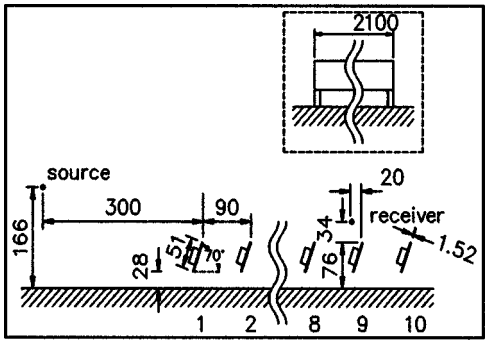


FIG. 1. Side elevation of geometry for measurements (insert is front elevation). Dimensions in mm for 1:10 scale model.

auditorium. The results reported are also only for unoccupied seating. There is evidence that the presence of an audience has little effect on the attenuation[1,2,12] of *absorbent* seats and so it is thought that the conclusions drawn will apply to the majority of occupied auditoria.

A. Scale model measurements

The 1:10 scale model seats used here had previously been developed to accurately represent the absorption and shape of typical unoccupied seats found in auditoria.[13] Measurements were carried out in both a semi-anechoic chamber and a large reflection-free zone. In the latter, time gating of the impulse response was used to remove the effect of the reflections from the walls. A typical measurement arrangement is shown in Fig. 1. The source was a 2.5-cm tweeter and the sound field was sampled by a standard $\frac{1}{4}$-in. microphone. Measurements were carried out using an unmodified commercially available maximum length sequence measurement system.[14] All frequency responses and impulse responses have been normalized (by deconvolution) to an anechoic calibration of the measurement system. The anechoic response of the measuring chain is shown in Fig. 2 along with a normalized version. The normalized response is a good approximation to a pure impulse and this shows that the deconvolution introduces few errors. All the results are shown as if for a full-scale model.

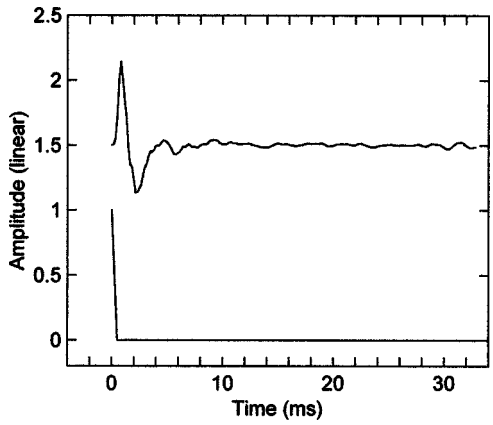


FIG. 2. Anechoic impulse response of the scale model measurement system (top) and the same response normalized (bottom).

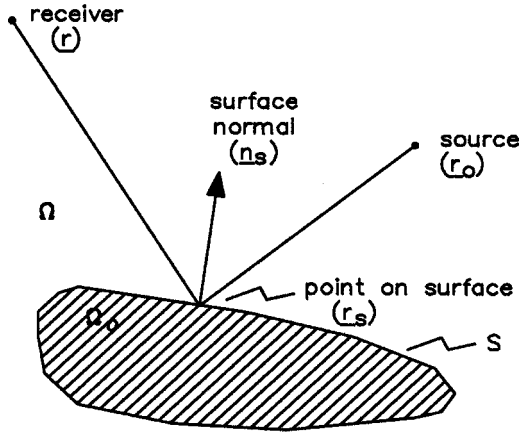


FIG. 3. Geometry for the Helmholtz–Kirchoff integral equation.

B. Boundary element method

The scattering from the seats has been predicted using a two-dimensional BEM.[15] The boundary element formation was based on the single frequency form of the Helmholtz–Kirchhoff integral equation for completely rigid surfaces. In this case the pressure $P(\underline{r})$ for one point source was

$$\int_S P(\underline{r}_s)\nabla G(\underline{r},\underline{r}_s).\underline{n}_s(\underline{r}_s)\,dS + P_i(\underline{r},\underline{r}_0) = \begin{cases} P(\underline{r}), & \underline{r}\in\Omega, \\ \frac{1}{2}P(\underline{r}), & \underline{r}\in S, \\ 0, & \underline{r}\in\Omega_0, \end{cases} \tag{1}$$

where $P_i(\underline{r},\underline{r}_0)$ was the sound pressure direct from the source; $\underline{n}_s(\underline{r}_s)$ was the outward pointing unit vector normal to the surface at $\underline{r}_s$, and $G(\underline{r},\underline{r}_s)$ was the Green's function. The Green's function was the standard two-dimensional form:

$$G(\underline{r},\underline{r}_s) = -\frac{i}{4}[H_0^{(1)}(k|\underline{r}-\underline{r}_s|) + H_0^{(1)}(k|\underline{r}-\underline{r}_s'|)], \tag{2}$$

where $H_0^1(x)$ was the Hankel function of the first kind of order zero. The second term in Eq. (2) was used when dealing with half space and creates image source effects, $\underline{r}_s'$ being the location of the image source. Figure 3 shows definitions of the vectors used.

The BEM solution technique involved first subdividing the surface into a set of elements across which the pressure was assumed constant. For this a subdivision of $\lambda/16$ or smaller was used. This small subdivision was required because the interference pattern close to the seating array was delicate and sensitive to changing conditions. Once the surface was subdivided the calculation proceeded in two steps: first an evaluation of the surface pressures was made via simultaneous equations, then the pressures at external receiver positions were calculated by a simple surface integral. The CHIEF[16] was used confirm unique solutions. No allowances for corners and edges were made in the application of Eq. (1). Two-dimensional methods were used for the predictions, with a cross section through the seating array being

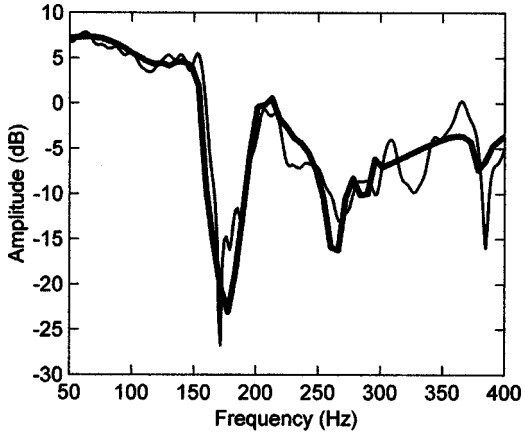


FIG. 4. Comparison of seat dip spectrum measured (light) above a set of hard seats with BEM prediction (bold).

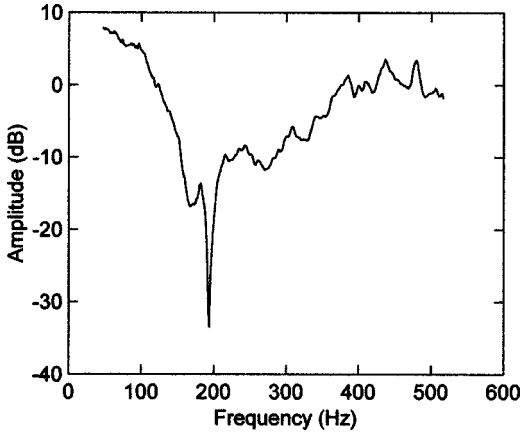


FIG. 5. A typical measured seat dip spectrum (absorbent seats).

defined; this vastly reduced the number of surface elements compared to three-dimensional BEMs and thus greatly decreased calculation times.

To test the validity of the prediction method, the pressure above a simple array of hard seats was compared to scale model measurements. The geometry modeled is as shown in Fig. 1, except that the absorbent parts of the seats were removed. This gave a stricter test, with sharply changing interference patterns close to the seats. Figure 4 shows a comparison between the experimental results and the BEM predictions. The spectrum is shown at a typical seat, where the sound had passed over eight rows of seats before reaching the microphone. Figure 4 shows that a very good degree of agreement had been achieved, especially considering this was a delicate interference pattern between many reflections, with the source and receiver very close to the seating array.

III. RESULTS AND DISCUSSION

A. Temporal and spectral features

The first measurements of seat dip attenuation in the literature were made using tone-bursts and single-cycle sinusoids.[1,2] More recent investigations have typically used the method employed here: measure a short impulse response propagated across seating and Fourier transform it to reveal a spectrum showing a seat dip. While the dramatic attenuation due to the seat dip effect is usually arrestingly clear in the frequency spectrum, there are important clues in the time domain that point the way to a clear understanding of the effect and suggest ways of controlling it.

Figure 5 shows a typical seat dip spectrum. It was measured using ten rows of absorbent scale model seats, as shown in Fig. 1. The microphone was in row 8, so that the spectrum in Fig. 5 is of sound passing over seven rows of seats. The size, frequency and shape of the attenuation are not untypical of those measured in auditoria. Figure 6 shows the corresponding impulse response: of particular interest are the strong reflections in the first 4 ms. There is a large negative reflection immediately after the direct sound which seems to come from the top of the seat in front of the microphone. This is followed by three quite strong positive reflections from the seat and floor close to the microphone. Many smaller reflections arrive at the microphone after this, up to and beyond the arrival time of the first reflection from the walls or ceiling of any real auditorium. If significant amounts of sound energy arrive so early from the seats, does the seat dip attenuation also appear early? This question is answered by changing the window applied to the impulse response before using the fast Fourier transform (FFT). To plot the evolution of the seat dip spectrum, one starts with a very narrow window containing only the direct sound, and calculates the spectrum. The start of the window is held just before the direct sound arrival, while the end is moved forward in small steps, taking a FFT at each increment. This was done to the impulse response in Fig. 6 and the results plotted as a surface in Fig. 7. What is perhaps most interesting about Fig. 7 is that the dip is established very early indeed. At 4 ms (full scale) after the direct sound, the dip appears. Thereafter its frequency changes little; the attenuation changes (mainly deepens) with the arrival of successive reflections from the seats and floor. This suggests that to control seat dip attenuation, one should pay the most attention to the earliest reflections from the seats and floor. By changing the amplitude and/or arrival time of these reflections, it should be possible to affect the interference pattern that produces seat dip attenuation.

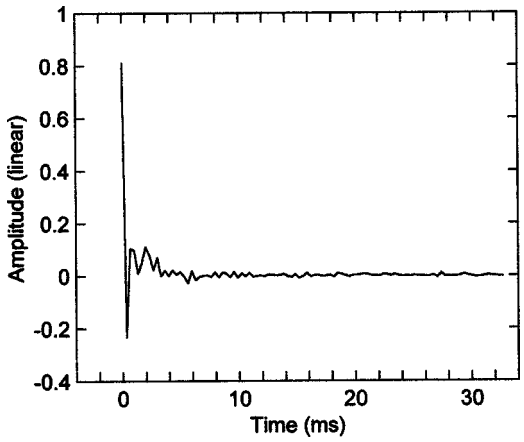


FIG. 6. A typical measured seat dip impulse response, used to produce Fig. 5.

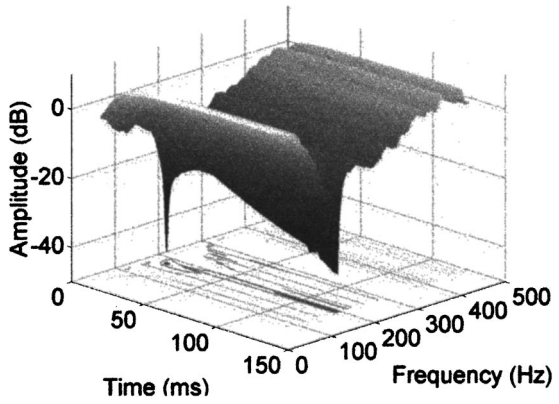


FIG. 7. Evolution of a typical seat dip spectrum over the period immediately after the direct sound arrives.

This can be simply tested in a rather crude fashion by deleting a reflection from the impulse response and viewing the resultant spectrum. In Fig. 8 this has been done, first for the initial negative reflection from the seat top, and, second, for the subsequent two reflections from the seat/floor. Figure 9 shows that, in both cases, deleting the reflections reduces the attenuation minimum by about 20 dB. The rest of the spectrum is largely unaltered. This initial success prompted attempts to control these reflections by modifying a modelled seat geometry.

B. Subjective threshold

In order to evaluate the effect of the different schemes described here, the predicted attenuation is compared with the subjective threshold. In experiments reported elsewhere, the subjective threshold for perception of seat dip attenuation has been measured.[11] The method used ten trained subjects in a realistic simulation of a concert hall sound field. A seat dip filter was applied to several of the early reflections and this resulted in an impulse response in which the seat dip attenuation varied over time, as it does in a real hall. The attenuation of the filter could be changed to find the attenuation that could be just detected by half of the subjects. There are several possible ways of expressing this threshold. In Ref. 11, the energy in the impulse response in the 200-Hz

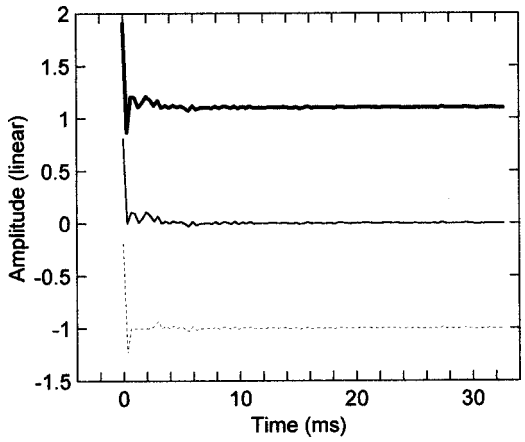


FIG. 8. The impulse response from Fig. 6 (bold) with the seat top reflection removed (light) and the early floor reflections removed (dashed).

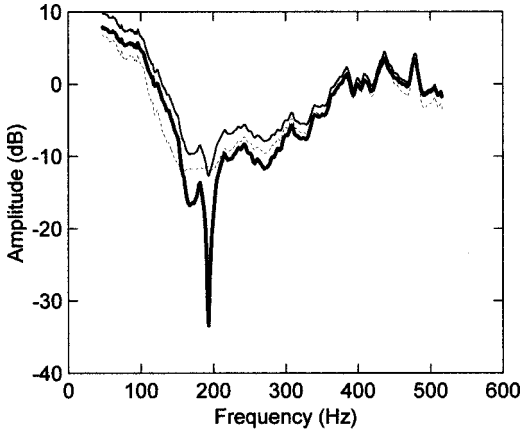


FIG. 9. The seat dip spectrum from Fig. 5 (bold) with the seat top reflection removed (light) and the early floor reflections removed (dashed).

octave band from 0 to 80 ms was used. This integration time was used because it covers all the reflections attenuated and also because 80 ms is commonly used to calculate musical clarity. When expressed like this, the threshold is a change of -3.8 ± 0.2 dB in the 200-Hz octave band early energy, from 0 to 80 ms.

To use the subjective threshold here, however, it must be modified. The predictions and measurements reported in this article included no reverberation and no early reflections other than those from the seats and floor. The threshold must therefore be described by the attenuation of the direct sound only. Thus the integration limit was changed from 80 to 18 ms and the impulse responses from the simulated concert hall reanalyzed. This revealed that the threshold would be -5.9 ± 0.4 dB octave band attenuation, from 0 to 18 ms. This means that, if the direct sound in a measured impulse response shows an attenuation of greater than 5.9 dB in the 200-Hz octave band, then 50% of listeners should be able to detect it (though they would base their perception on the total early sound field). It should be noted that the shorter integration limit has produced a larger threshold attenuation. This is because there are always several reflections in an early sound field which are not subject to seat dip attenuation. These arrive after the direct sound and have the effect of masking the subjective perception of the effect by increasing the low-frequency early energy. Thus, if one takes the impulse response at the subjective threshold, the attenuation is larger at 18 ms than it is at 80 ms.

Using the revised subjective threshold involves making an assumption about the early sound field in the concert hall one is considering: that it is similar to the early field in the simulator. As reported in Ref. 11, the simulation was designed to be representative of a ''generic'' good shoebox hall: values of parameters measured in it are given in that paper. Clearly, different halls have different early fields and so the relationship between the attenuation in the direct sound and the attenuation in the total early field may vary from one hall to another. For example, a hall with ceiling reflectors will tend to have a smaller attenuation in the total early field than one without. This means that hall designers should use the more general 80 ms early energy version of the subjective threshold where possible: the -5.9-dB value

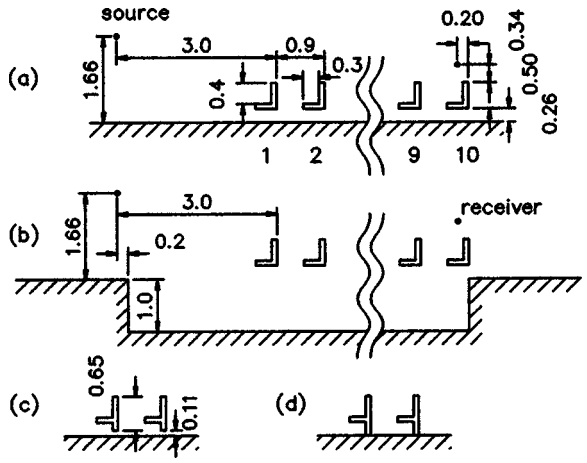


FIG. 10. Seat profiles used in BEM predictions: (a) standard shape, (b) with floor pit, (c) with small underpass, and (d) with no underpass. Dimensions in meters for full scale.

should only be used where only the direct sound is available, as here. It should also be noted that differences in the reverberant field will have little effect on perception of seat dip attenuation.[11]

C. BEM predictions for modified seats

Many different geometries, some with absorbing and diffusing elements, were modeled. Figure 10 illustrates some of the basic shapes. Figure 10(a) is the base configuration. A seat shape was adopted which allowed a realistic model of the effect of different absorbent elements on the seats. The surface impedance of each element on the seat was taken from impedance tube measurements of parts of a typical well-upholstered auditorium seat. Data was available for both the seat squab (78-mm polyurethane foam) and backrest (23-mm polyurethane foam). In the base configuration, the back of the backrest and the underside of the squab had zero admittance. In other configurations, the effects were sought of altering the admittance of different parts of the seats.

In one respect, the BEM used here is not complete. It does not model any transmission path through an absorbent seat cushion and out the other side. However, most elements of a real seat consist of absorbent cushion backed by a rigid surface, so that any transmitted levels are likely to be very low. Transmission through the seat is therefore assumed to be not significant in the total signal at the receiver.

Figure 10(b) illustrates one of the many geometries investigated empirically to change the interference pattern by altering the timing of the seat or floor reflections. In this profile, a 1-m pit is introduced under the seats. Several other floor profiles were tried, some including diffusing and absorbing elements. Finally, Figs. 10(c) and (d) show the underpass modifications explored.

1. Basic configuration

The magnitude and frequency of seat dip attenuation are known to vary with the number of seat rows in front of the receiver (r), the receiver height, and the vertical (θ) and horizontal angles of incidence of the direct sound.[2] The results reported here have a fixed source position so that r and θ are

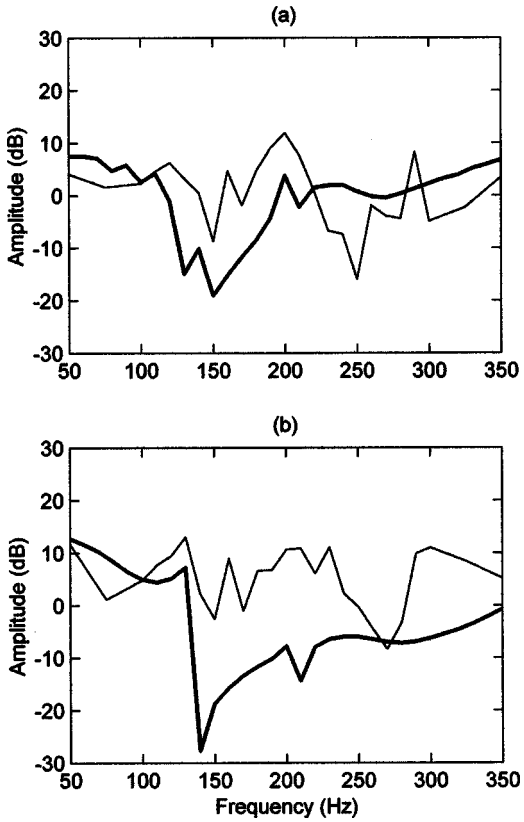


FIG. 11. Effect of floor pit on predicted spectrum at (a) row 3 and (b) row 8: standard seats (bold), with pit (light).

increasing together. Using the standard configuration of Fig. 10(a) with standard admittances, the octave attenuation is less than -5.9 dB for values of r from 1 to 5. For r from 6 to 9 (and beyond) the attenuation gets steadily worse. Hence, seat dip attenuation would be audible at all rows further from the stage than $r=5$ and remedial treatment is necessary at these seats.

2. Floor pit

For $r>5$, the pit is the best method. Figure 11 shows two typical results for an empty 1-m pit. At $r=8$, the introduction of the pit reduces the octave attenuation from -11.1 dB to $+3.1$ dB. Several variants on the basic empty pit were also tried—sloped floors, diffusing lining, subdivided pits—but did not produce any significant benefit. It was found that the pit needed to be large and span many seating rows. It is envisaged that the pit could be provided by using a grid, which would be acoustically transparent at low frequencies and covered with carpeting to conceal the acoustical treatment. It might be possible that the pit space could be used for ventilation systems, although care would have to be taken to ensure that this would not create background noise problems. It certainly would be possible to use the space for other services such as electricity. This solution to the seat dip effect was found to be quite robust, the most important feature being the requirement for the pit to be open to many rows of seats. For example, simulation of cross beams to support seating—resulting in smaller entrances to the pits in each seating row—did not affect the decrease in the dip.

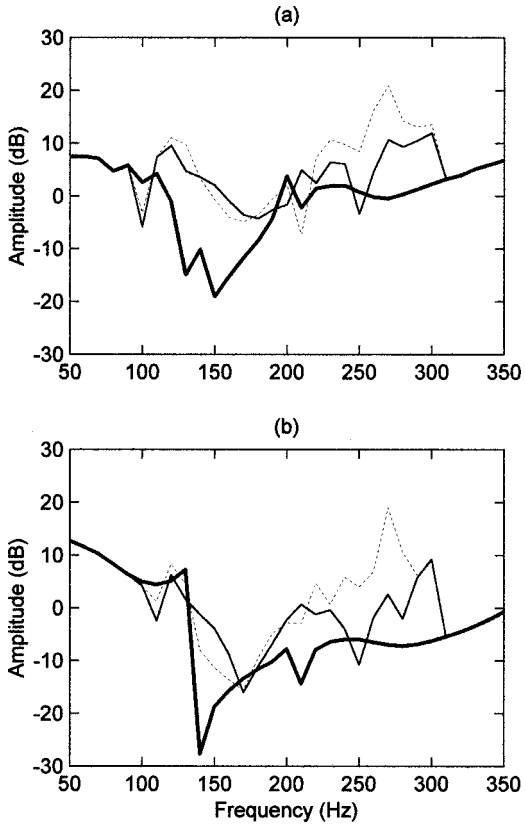


FIG. 12. Effect of seat absorption on predicted spectrum at (a) row 3 and (b) row 8: standard seats (bold), with resonantly absorbing squab underside (light), with resonantly absorbing squab underside and seat top (dashed).

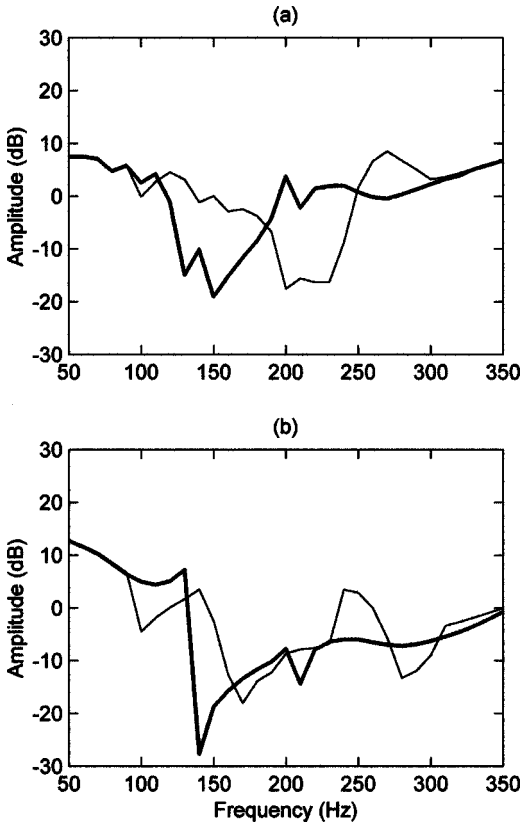


FIG. 13. Effect of floor absorption on predicted spectrum at (a) row 3 and (b) row 8: standard seats (bold), with resonantly absorbing floor (light).

3. Changes in seat absorption

Next, attempts were made to control early reflections from the seats and floor with absorption. This was done by changing the admittance of some elements of the seat or floor to give an absorption coefficient of 1 at frequencies from 100 to 300 Hz. The most extreme absorption was used to give the best chance of seeing an effect. A real implementation would be likely to have a lower absorption. Changing the underside of the seat squab from a hard surface to a resonant absorber can control sound passing through the seat underpass. This has a significant effect on the attenuation at the receiver, as Fig. 12 shows. The audible dip at $r=8$ (-11.1 dB octave attenuation) has been moved above the subjective threshold to -4.7 dB. The reduction in attenuation is more noticeable at low values of r; when the attenuation becomes larger, the improvement is sometimes not great enough to render the dip inaudible. It is thought that a resonant seat squab could be achieved in a real seat by drilling the squab underside to form Helmholtz resonators or by making the underside flexible enough to behave as a panel absorber. It is, of course, possible that doing this would increase the random incidence absorption in the audience area. Any increase might be small when the seats are occupied, however, because the audience would shield the resonators from the main reverberant field.

The same technique was also used to try to control the large negative reflection from the seat tops by making these highly absorbent also. As well as being much harder to achieve practically, this has little *additional* effect, as Fig. 12 shows. A combination of resonant squab and floor pit was also tried—this offered a slight improvement over the pit alone at some seats and none at others. The spectrum also became less flat in general. It seems likely that an audience would have an effect on the performance of the resonantly absorbing seat top. Because the improvement seen here was small, however, the effect of an audience was not pursued.

Finally, the floor under the seats was made resonantly absorbing, modeling a scheme suggested by Ando *et al.*[8] [In order that the Green's function of Eq. (2) could be used here, the floor absorbers were modeled as absorbent blocks placed a small distance above a hard floor.] The benefits of floor absorbers were less ''well-behaved'' than those of seat squab absorbers, as Fig. 13 shows. Floor absorbers help at some seats, for example at $r=8$ in Fig. 13(b). At other seats, however, the dip is changed only in frequency, as happens at $r=3$ in Fig. 13(a). These results agree only partially with Ando's original predictions, where floor absorption is shown to reduce the dip with no other effects. The mixed results in Fig. 13 agree rather better with measured data on floor absorbers in a real hall.[4] It seems that if the introduction of extra absorption into the audience area is permissible, then it will have a greater beneficial effect on the seat squabs rather than on the floor. This may be because the seat squabs are closer to the point of maximum particle velocity and are therefore more effective: at 200 Hz, the quarter-wavelength will be 43 cm from the floor for normal incidence, and less for oblique incidence.

It is interesting to compare the performance of floor absorbers with that of the floor pit. When the pit is used, the

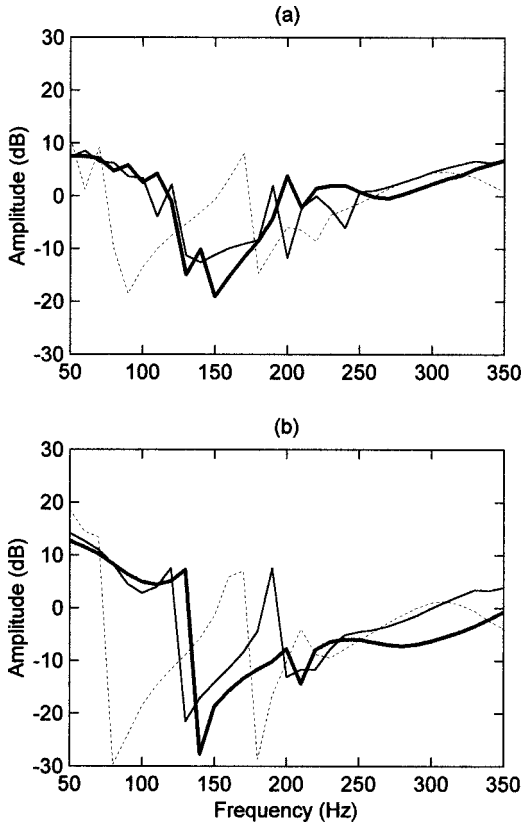


FIG. 14. Effect of seat underpass on predicted spectrum at (a) row 3 and (b) row 8: full underpass (bold), half underpass (light), and no underpass (dashed).

arrival of sound from the floor is altered in time and phase. When floor absorbers are introduced, the sound scattered from the floor is strongly attenuated. The floor absorbers might be seen as similar to an infinite pit. It was found empirically when testing the pit that 1 m was an optimum size. Larger and smaller pits did not reduce the attenuation as well, or did so less consistently. This is because the scattered sound from the floor is combined at the receiver with sound scattered from the seats. The attenuation is reduced more reliably by optimizing the way in which these different reflections combine instead of removing just one. The only alternative would seem to be to remove all the sound scattered from the seats and floor to leave an anechoic direct sound, but this is less practical.

4. Seat shape

As well as altering the admittance of the seat surfaces, changes in the shape of the seats were investigated. Many profiles were investigated, though all were constrained to bear a passing resemblance to a shape that a listener could sit on! It was found, perhaps surprisingly, that the basic shape adopted in Fig. 10(a) could not be improved upon. Some profile changes made the attenuation worse, while others made little difference. Restricting the underpass as in Figs. 10(c) and (d) makes the attenuation worse at all seat positions. Figure 14 shows that it is quite possible to make an inaudible dip into an audible one, though a half-underpass is much better than no underpass at all. Changing the shape of the squab underside, using several scattering profiles, had little effect. This is mainly because considerable space is needed to introduce features which will affect the wavelengths of interest.

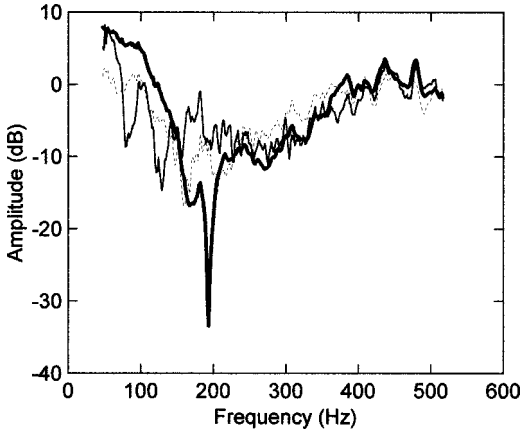


FIG. 15. Effect of floor pit on measured spectrum at row 7: absorbent seats without pit (bold), with pit (light), with pit filled with foam (dashed).

D. Scale model confirmation of reduced attenuation

Once the pit method had been identified as reducing the dip, it was decided to test the arrangement in 1:10 model scale measurements. Ten rows of absorbent seats were used, with the hard floor measurements arranged as in Fig. 1. Measurements were also made on the same seats with a 1.1-m pit (full scale size) beneath them. A wire mesh and cotton sheet material were used to model the effects of grill and carpet expected in the full scale system. (The mesh and sheet were found, in fact, to have little influence on the attenuation.) The pit was tested empty and partially filled with polyurethane foam (to further attenuate floor reflections). It should be noted that the profile of the scale model seats is not identical to the BEM ones, for reasons of practical expediency, so exact agreement between results is not expected. However, the two sets of seats are expected to show the same trends.

Figures 15 and 16 show comparisons of the sound fields with and without the pit for two typical seats. It can be seen that the dip has been significantly reduced, but has not been completely removed. At all measuring positions where the

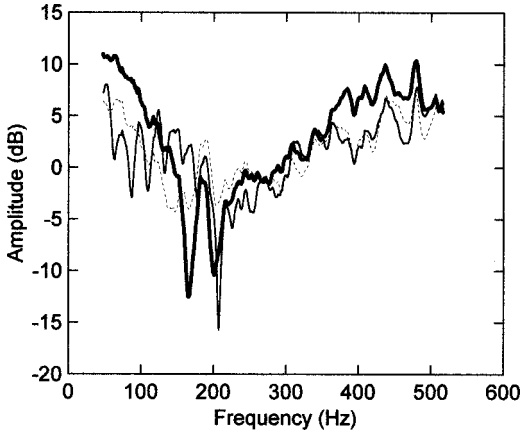


FIG. 16. Effect of floor pit on measured spectrum at row 3: absorbent seats without pit (bold), with pit (light), with pit filled with foam (dashed).

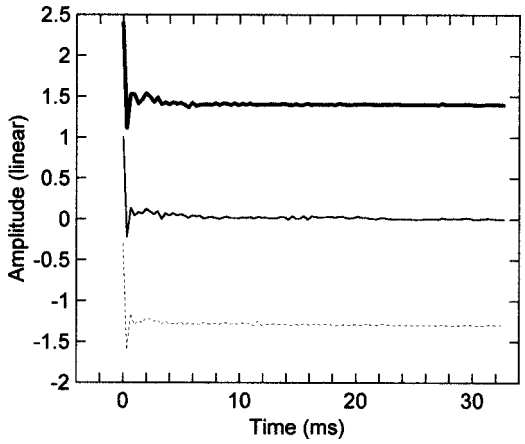


FIG. 17. Effect of floor pit on measured impulse response at row 7: absorbent seats without pit (bold), with pit (light), with pit filled with foam (dashed).

dip would have been subjectively significant, the pit reduced the octave band attenuation to below the threshold of perception. The use of foam in the pit did not significantly improve the frequency responses; while the foam lessened the dip at some seats, it increased the dip at others. The foam does seem to deal with the comb filtering at very low frequencies, which is introduced by the pit. This extra interference effect is unlikely to be subjectively significant, though. Using foam in the pit would tend to increase the random incidence absorption of the audience and so is not a favored solution.

Finally, Fig. 17 shows that the floor pit has a measurable effect on the early seat/floor reflections in the impulse response. Introducing foam into the pit further reduces the reflections, as expected, but the effects in the frequency domain are not so significant.

IV. CONCLUSION

This article has explored the options available for controlling seat dip attenuation in concert halls. The aim was to bring the level of the attenuation below the subjective threshold. Examining the effect in the time domain showed that the dip is introduced within 4 ms of the direct sound arrival by sound scattered from the seats and floor. Controlling this scattered sound gave the desired reduction in attenuation. Several changes to the profile of the seats and floor were tested using a boundary element model. The most successful change was the introduction of a pit under the seats: this consistently reduced the attenuation to below the subjective threshold. Scale model measurements were used to confirm the size of the improvement. The next most effective method was to make the underside of the seat squab highly absorbing at the main dip frequency. Other methods, such as resonant floor absorbers, were less effective.

It is likely that no single solution to the seat dip effect will be appropriate for every auditorium. Different halls will have different constraints on the shape of the seats and floor. However, where there is scope for changes to the shape or impedance of these elements, this should be explored. Hall designers are already used to controlling the sequence of early reflections from the walls and ceiling of an auditorium. The results presented here suggest that they should also pay attention to sound arriving at the listener between the direct sound and the first major early reflection.

ACKNOWLEDGMENT

The authors are very grateful to Tim South for constructing the scale model and assisting in measuring it.

[1] G. M. Sessler and J. E. West, ''Sound transmission over theater seats,'' J. Acoust. Soc. Am. **36**, 1725–1732 (1964).
[2] T. J. Schultz and B. G. Watters, ''Propagation of sound across audience seating,'' J. Acoust. Soc. Am. **36**, 885–896 (1964).
[3] J. S. Bradley, ''Some further investigations of the seat dip effect,'' J. Acoust. Soc. Am. **90**, 324–333 (1991).
[4] W. J. Davies and Y. W. Lam, ''New attributes of seat dip attenuation,'' Appl. Acoust. **41**, 1–23 (1994).
[5] K. Ishida, K. Sugino, and I. Masuda, ''On the sound reflection of the auditorium seats,'' Proc. 13th I.C.A., Belgrade (1989), Vol. 2, pp. 157–160.
[6] E. Mommertz, ''Some measurements of the propagation of acoustic-waves skimming over the public and seats,'' Acustica **79**, 42–52 (1993).
[7] Y. Kawai and T. Terai, ''Calculation of sound fields over audience seats by using integral method,'' Trans. ASME, J. Vib. Acoust. **113**, 22–27 (1991).
[8] Y. Ando, M. Takaishi, and K. Tada, ''Calculations of the sound transmission over theater seats and methods for its improvement in the low-frequency range,'' J. Acoust. Soc. Am. **72**, 443–448 (1982).
[9] J. LoVetri, D. Mardare, and G. Soulodre, ''Modelling of the seat dip effect using the finite-difference time-domain method,'' J. Acoust. Soc. Am. **100**, 2204–2212 (1996).
[10] D. Takahashi, ''Seat dip effect: The phenomenon and the mechanism,'' J. Acoust. Soc. Am. **102**, 1326–1334 (1997).
[11] W. J. Davies, T. J. Cox, and Y. W. Lam, ''Subjective perception of seat dip attenuation,'' Acustica **82**, 784–792 (1996).
[12] J. S. Bradley, ''On the effect of the audience on auditorium acoustics,'' Proc. 15th I.C.A., Trondheim (1995), Vol. 2, pp. 389–392.
[13] W. J. Davies, ''The effects of seating on the acoustics of auditoria,'' Ph.D. thesis, University of Salford, UK, 1992.
[14] D. D. Rife and J. Vanderkooy, ''Transfer-function measurement with maximum-length sequences,'' J. Audio Eng. Soc. **37**, 419–444 (1987).
[15] T. Terai, ''On the calculation of fields around three-dimensional objects by integral equation methods,'' J. Sound Vib. **69**, 71–100 (1980).
[16] H. A. Schenek, ''Improved integral formulation for acoustic radiation problems,'' J. Acoust. Soc. Am. **44**, 41–58 (1968).

Image shift caused by strong lateral reflections, and its relation to inter-aural cross correlation

Toshiyuki Okano
Graduate School of Science and Technology, Kobe University, Rokkodai, Nada, Kobe, 657-8501 Japan and Takenaka Research and Development Institute, 1-5-1, Ohtsuka, Inzai, Chiba, 270-1395 Japan

(Received 25 January 2000; revised 18 July 2000; accepted 24 July 2000)

Conditions of sound fields were studied to avoid the occurrence of image shift, an acoustical phenomenon caused by extremely strong lateral reflections. Three types of image shift are defined as variations of the normal condition, in which there is only one auditory image perceived in the source direction. From a series of listening tests in electrically reproduced sound fields, each of the three types of image shift was found to have its own specific frequency range in which the image shift occurred. They were also found to relate to each of three measures which are all derived from the inter-aural cross correlation function (IACF) of binaural impulse responses of a sound field. The results of measurement made in three existing concert halls suggested acceptable ranges for the measures, in which the image in the source direction is not lost in the 500 Hz and higher frequency bands. However, the 250 Hz and lower frequency bands were found not so important in evaluating sound fields of concert halls in relation to the image shift. Finally, 0.85 is found to be the upper limit for the preferable range of $[1-\mathrm{IACC}_{E3}]$.
[S0001-4966(00)00411-2]

PACS numbers: 43.55.Fw, 43.55.Gx, 43.55.Hy [JDQ]

I. INTRODUCTION

Apparent source width (ASW) has been found to be created primarily by lateral reflections and is one of the important attributes of the image source in evaluating the acoustical quality of concert halls. Okano *et al.* (1998) showed that high ranking halls in rating show large values for both $[1-\mathrm{IACC}_{E3}]$ and G_{low}, which are both defined as combined measures of ASW, where IACC_{E3} is defined as an arithmetic average of inter-aural cross-correlation coefficients derived from the first 80 ms of impulse response after the arrival of the direct sound in three 1/1 octave bands with mid-frequencies of 500, 1k, and 2k Hz (Hidaka *et al.*, 1995), and G_{low} is the strength in the lower frequency range below 355 Hz. For highest ranking halls among them, $[1-\mathrm{IACC}_{E3}]$ exceeds 0.7 in hall average value. They sometimes reach 0.8 or more for some individual seats. These large $[1-\mathrm{IACC}_{E3}]$ values arise from strong lateral reflections, of which angles of incidence are widely distributed. Here arises a question whether or not a very large $[1-\mathrm{IACC}_{E3}]$ of about 0.9 or more will produce better conditions for a concert hall. Such a kind of sound field would be realized by a very large numbers of strong lateral reflections, that is, as many as 10 or more in the first 80 ms of impulse response, or possibly by electrical enhancement of lateral reflections, for instance, with loudspeakers on side walls.

In an earlier study, Barron (1971) pointed out that a lateral reflection having too large an amplitude will create image shift. This phenomenon, described as movement of the source image from the direction of direct sound toward the direction of a lateral reflection, obviously produces an undesirable element to the acoustical quality of a hall. In this case, strong lateral reflections produce unfavorable conditions rather than enlarging the ASW. However, it has not been studied further how strong lateral reflections are when they cause image shift and how to judge the occurrence. Some lateralization studies tried to predict the image position from binaural signals, but these attempts are not intended to judge the occurrence of image shift (Bilsen *et al.*, 1973; Stern *et al.*, 1978, 1988; Shackleton *et al.*, 1992).

The present study intends to determine the status of image shift affected by strong early lateral reflections and to find measures to judge if a sound field would produce image shift or not. This subject is of particular importance in designing a concert hall that should aim at winning the highest class reputation, i.e., $[\mathrm{A}^{+},\mathrm{A}]$ by Beranek (1996). To be specific, a hall should have its $[1-\mathrm{IACC}_{E3}]$ to be as high as 0.8 or more for some individual seats. Also, this subject would be of interest to readers who work for auralization of a sound field in a room.

II. DEFINITION OF IMAGE SHIFT AND IMAGE SHIFT CATEGORIES

Barron (1993) described the image shift as ''...if large plane surfaces are used... the sound of a particular instrument suddenly appears to come from the reflector instead of from the stage. This can occur, for instance, when trumpets playing in their upper register (where it is highly directional), point at such a reflector.'' Image shift is also sometimes observed in opera houses when the orchestra is playing in the orchestral pit. If the floor level of the orchestral pit is too low, the orchestral sound can be localized in the direction of surfaces that provide early reflections. This occurs because the direct sound is diffracted too much by the balustrade and has considerably smaller amplitude than early reflections. From this observation, in the present study, the image shift is defined as a disagreement in the directions between the real

source and the auditory image, which is affected by the large amplitude of lateral reflections relative to that of the direct sound.

It is supposed that the movement of an image from the source direction toward lateral directions is not a sudden jump. A series of lateralization studies suggest the directional movement of a sound image from the direct sound to a lateral reflection is continuous. The location of the sound image is distributed continuously from the right ear to left according to the condition of binaural time and level difference (Yost *et al.*, 1987, pp. 49–52). Furthermore, it is not always a single image that is perceived in a certain direction. Bilsen *et al.* (1973) reported that there are cases of perceiving two sound images simultaneously at different positions for a frequency range above 1 kHz. To make sure, this is quite different from ''echo'' because ''echo'' is a phenomenon of perceiving more than two images separately in the time domain. Another possibility of perceiving plural images simultaneously is when a pair of sounds is presented to each ear in opposite phase. In this case, two sound images are perceived separately at a position corresponding to each ear (Yost *et al.*, 1987, p. 53; Damaske *et al.*, 1972). These experiences imply possibilities of perceiving more than two simultaneous images for one source in the sound field in a room. Finally, these kinds of possibilities for source images are categorized into one normal image category and three image shift categories as below.

Category 0: One image is perceived and the source direction is included within the image.
Category 1: One or more lateral separate images are perceived in addition to the image in the source direction.
Category 2: Two separate lateral images are perceived, each of which does not include the source direction.
Category 3: One lateral image is perceived, which does not include the source direction.

Category 0 is the normal image category naturally experienced in usual rooms. In Category 1, the image is ''shifting'' away from the source direction, although the direction of the real source and either of the auditory images are still in agreement. In the other two categories, the image in the source direction is lost, resulting in disagreement in the direction between the real source and the auditory images. Conceptual drawings for these categories are shown in the following section (Fig. 2). It is considered that the larger the category number, the greater the image shift. Note that Category 3 is the case of the most apparent image shift.

Generally, two sound field attributes are considered to contribute to generation of the image shift: one is an unsymmetrical incidence of reflections in energy and their angles of incidence and the other is the existence of strong lateral reflections. The influences of unsymmetrical incidence on the perceived source direction has been examined in past lateralization studies, such as Stern *et al.* (1978, 1988) and Shackleton *et al.* (1991). They experimented on binaural stimuli with varying inter-aural level differences and inter-aural time differences and learned that the peak position of inter-aural cross-correlation function, defined in their auditory model, indicates the direction of the sound source perceived. These results agree with the discussion made by Damaske *et al.* (1972) and Ando (1977) that the peak in the IACF occurs in the delay that corresponds to the source direction. The present study, however, focuses on the other factor, the existence of strong lateral reflections.

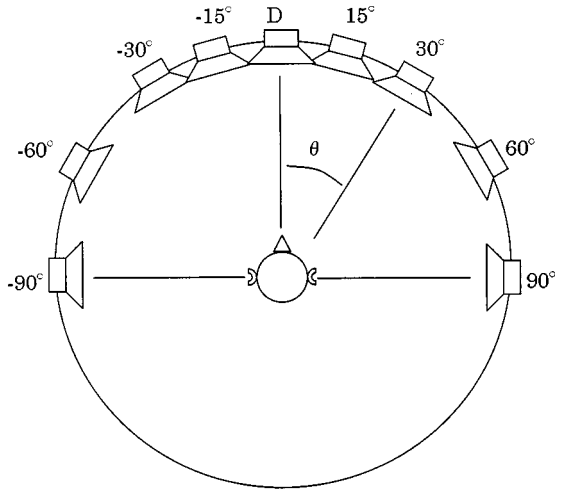


FIG. 1. Arrangement of multiple loudspeakers used in the listening test. Loudspeaker D produces the direct sound, to which the subject always faces. Other loudspeakers emit early lateral reflections with different time delays and angles of incidence.

III. EXPERIMENT

The listening test was conducted to determine the categories of the source image in various types of sound fields composed of direct sound and strong early lateral reflections.

A. Procedure of listening test

Sound fields were electrically reproduced in an anechoic chamber with a multiple loudspeaker system (Fig. 1). The loudspeaker just in front of the subject supplies the direct sound and the others supply early lateral reflections. A mu-

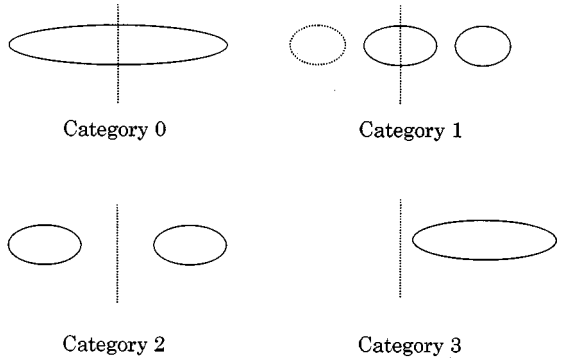


FIG. 2. Normal image category and three image shift categories. Category 0: normal; Category 1: addition of lateral separate image(s); Category 2: two separate lateral images; Category 3: one lateral image. Vertical dotted line indicates the real source direction.

sical signal, taken from the first movement of Bruckner's Symphony No. 4, was fed into the reproduction system through 1/1 octave band filter with a center frequency ranged from 125 to 4 k Hz. Since the image shift is considered to occur in higher frequencies and show different aspects in different frequency ranges, the 1/1 octave band filtered signal was used to determine effective frequency bands for the image shift.

The musical signal was reproduced repeatedly until every subject feels sure about his perception of the image shift. Then each subject wrote down the shape of the perceived sound image on an answer sheet with a scale along the loudspeaker positions printed thereon in advance. The results were classified into four categories: the normal one and three previously defined ones. Conceptual drawings for these categories are shown in Fig. 2.

Four male subjects participated in the experiment. Three of them were acoustical engineers and a fourth a student. Each subject judged the same test condition twice. As a result, eight answer sheets were checked for each of six 1/1 octave bands of each test condition.

B. Test sound fields

Forty-three types of sound fields were designed, with the number, initial delay, the angles of incidence, and level of early reflections relative to the direct sound used as parameters. Composition of each test sound field is shown in Table I.

The number of early lateral reflections was two, four, and eight. In each test sound field, reflections come in a pair, sharing the same amplitude but reaching from symmetrical directions with the direct sound direction in the center. For the cases with two reflections, the sound was caused to have a varying incident angle (θ) of $\pm 15°$, $\pm 30°$, $\pm 60°$, and $\pm 90°$. For the cases with four reflections, θ was fixed at $\pm 30°$ and $\pm 60°$. For the cases with eight reflections, θ was then changed to be $\pm 15°$, $\pm 30°$, $\pm 60°$, and $\pm 90°$.

All the reflections had smaller delays than 80 ms to the direct sound. It should be noted that the range of delay smaller than 80 ms is known to be the effective element that causes source images and is therefore considered to be very critical in evaluation of a hall, as reported by Hidaka *et al.* (1995). For two reflections' cases, intentional change of initial delay was considered. The first setting was a combination of 16 and 24 ms for the paired reflections, and the other was a combination of 36 and 44 ms. The second combination was added for the cases with $\theta = \pm 30°$ and $\pm 60°$. For the cases of four reflections, the delay was set at 16, 24, 36, and 44 ms. The first two and the last two reflections were paired in their angles of incidence. For the cases with eight reflections, 16, 24, 30, 36, 44, 50, 60, and 74 ms of delay were selected. Every pair of symmetrical reflections in all cases had a relative delay difference of more than 8 ms. This is to avoid a kind of interference between the two symmetrical reflections, as reported by Okano *et al.* (1998).

The level of reflections was identified by L, which is defined by Eq. (1). L expresses the total sum of reflection energy in decibel relative to the direct sound:

$$L = 10 \times \log_{10} \frac{\Sigma A_n^2}{A_0^2}, \tag{1}$$

where A_n denotes the amplitude of the nth reflection and A_0 is the amplitude of the direct sound. From measurements of existing concert halls and theatres, L was found to be in the range approximately from 0 to 10 dB when a source is located on the stage and often exceeds 10 dB when a source is located in the orchestral pit, in the frequency bands of 500, 1k, and 2k Hz for the early 80 ms part of impulse responses. In the test sound field, L was selected among the range from 0 to 13 dB for the experiment. In order to control L, the level of the direct sound was decreased with the increase of the level of lateral reflections. The reproduction level was kept constant as long as the number of reflections remains the same. For the cases with four reflections, two different types of variation in the amplitude of reflections were considered to control L. One is that the level of reflections from $\theta = \pm 60°$ should vary, while that of reflections from $\theta = \pm 30°$ should be fixed constant. The other is to decrease the level of the direct sound and reflections from $\theta = \pm 30°$ and to increase the level of reflections from $\theta = \pm 60°$ as L increases. It was the same for the cases with eight reflections; the levels of reflections from $\theta = \pm 15°$ and $\pm 30°$ were kept unchanged for one series and were changed for the other case in the same manner.

The reproduction level of the musical stimuli was 70 dBA for the cases of two reflections in each frequency band signal; 73 and 75 dBA were used for the cases of four and eight reflections.

IV. IMAGE SHIFT CATEGORIES IN TEST SOUND FIELDS

Table II shows the results of the listening test. The perception frequency of image shift categories, from 0 to 3, is tabulated by the band in each test sound field, with figures under each category number showing the times of occurrence of image shift. For three image shift categories of 1, 2, and 3, figures equal to 3 or larger are bold. Among the cases in which the total frequency of the three image shift categories equal to 3 or larger, a frequency of 2 times is also bold when the difference from the highest frequency in each category is 1 or smaller. In Category 0, figures equal to 6 or larger are bolded. Among the bold figures in each frequency band of a sound field, the figure corresponding to the case having the largest number in its name was underlined. This underlined figure corresponds to the representative image shift category of the test sound fields, indicating the level of image shift occurrence the sound field reaches.

A. Dependence on the angle of incidence and delay

Figure 3 shows representative image shift categories for cases of two symmetrical lateral reflections. When an angle of incidence is $\pm 15°$, the image shift is not observed except for the case with L of 8 dB in a 4k Hz band. At 500 Hz and in the lower bands, $\pm 60°$ and $\pm 90°$ are sensitive to image shift occurrence. In the higher frequency bands, the sensitive

TABLE I. Test sound fields.

Name of test sound field	A11	A12	A13	A14	A21	A22	A23	A24	A31	A32	A33	A34	A41	A42	A43	A44	B21	B22	B23	B24	B31	B32	B33	B34
Number of reflections	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Total level of reflections *re* direct sound (dB)	8	5	3	0	8	5	3	0	8	5	3	0	8	5	3	0	8	5	3	0	8	5	3	0

Attributes of the direct sound and reflections

Angle of incidence (°)	Delay (ms)	Amplitude				Amplitude				Amplitude				Amplitude				Amplitude				Amplitude			
90	16													0.66	0.61	0.56	0.50								
60	16									0.66	0.61	0.56	0.50												
60	36																					0.66	0.61	0.56	0.50
30	16					0.66	0.61	0.56	0.50																
30	36																	0.66	0.61	0.56	0.50				
15	16	0.66	0.61	0.56	0.50																				
0	0	0.35	0.50	0.61	0.71	0.35	0.50	0.61	0.71	0.35	0.50	0.61	0.71	0.35	0.50	0.61	0.71	0.35	0.50	0.61	0.71	0.35	0.50	0.61	0.71
−15	24	0.66	0.61	0.56	0.50																				
−30	24					0.66	0.61	0.56	0.50																
−30	44																	0.66	0.61	0.56	0.50				
−60	24									0.66	0.61	0.56	0.50												
−60	44																					0.66	0.61	0.56	0.50
−90	24													0.66	0.61	0.56	0.50								

Name of test sound field	C11	C12	C13	C14	C15	C21	C22	C23	C24	D11	D12	D13	D14	D15	D21	D22	D23	D24	D25
Number of reflections	4	4	4	4	4	4	4	4	4	8	8	8	8	8	8	8	8	8	8
Total level of reflections *re* direct sound (dB)	12	8	5	2	0	8	7	6	5	13	10	7	5	3	10	9	8	7	5

Attributes of the direct sound and reflections

Angle of incidence (°)	Delay (ms)	Amplitude					Amplitude				Amplitude					Amplitude				
90	44															0.83	0.79	0.74	0.63	0.41
90	74										0.47	0.43	0.35	0.25	0.03					
60	16						0.93	0.87	0.78	0.56						0.83	0.79	0.74	0.63	0.41
60	36	0.65	0.60	0.49	0.34	0.03														
60	50										0.47	0.43	0.35	0.25	0.03					
30	16	0.71	0.71	0.71	0.71	0.71					0.71	0.71	0.71	0.71	0.71					
30	30															0.03	0.21	0.31	0.47	0.64
30	36						0.03	0.29	0.44	0.66										
15	36										0.71	0.71	0.71	0.71	0.71					
15	74															0.03	0.21	0.31	0.47	0.64
0	0	0.35	0.50	0.71	0.86	1.00	0.50	0.55	0.62	0.71	0.35	0.50	0.71	0.86	1.00	0.50	0.55	0.62	0.71	0.82
−15	44										0.71	0.71	0.71	0.71	0.71					
−15	60															0.03	0.21	0.31	0.47	0.64
−30	24	0.71	0.71	0.71	0.71	0.71					0.71	0.71	0.71	0.71	0.71					
−30	44						0.03	0.29	0.44	0.66										
−30	50															0.03	0.21	0.31	0.47	0.64
−60	24						0.93	0.87	0.78	0.56						0.83	0.79	0.74	0.63	0.41
−60	30										0.47	0.43	0.35	0.25	0.03					
−60	44	0.65	0.60	0.49	0.34	0.03														
−90	36															0.83	0.79	0.74	0.63	0.41
−90	60										0.47	0.43	0.35	0.25	0.03					

TABLE II. Occurrence frequency of image categories by the test sound field and frequency band. N: number of reflections, θ: angles of incidence, Δt_1: initial delay of the first reflection, L: total level of reflections relative to that of the direct sound. Among categories 1, 2, and 3, figures equal to 3 or larger are bolded. Among the cases in which the total frequency of the three image shift categories equal to 3 or larger, figures of 2 times is also bolded when the difference from the highest frequency in each category is 1 or smaller. In category 0, figures equal to 6 or larger are bolded. In each band of a sound field, underlined figure indicates correspondence to a representative category which was defined by the category having the largest number in its name among the bolded.

Name of sound field	Attributes of test sound field				1/1 oct. band center frequency																							
					125 Hz				250 Hz				500 Hz				1k Hz				2k Hz				4k Hz			
					Category				Category				Category				Category				Category				Category			
	N	Δt_1 (ms)	θ (°)	L (dB)	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
A11	2	16	±15	8	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**7**	1	0	0	**6**	2	0	0	5	**2**	1	0
A12				5	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**7**	1	0	0	**8**	0	0	0	**7**	1	0	0
A13				3	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0
A14				0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0
A21	2	16	±30	8	2	0	0	**6**	4	0	0	**4**	4	0	0	**4**	2	**4**	0	2	2	**6**	0	0	3	**3**	1	1
A22				5	**6**	0	0	2	**6**	0	0	2	5	1	1	1	3	**3**	0	**2**	3	**4**	0	1	4	**3**	0	1
A23				3	**7**	0	0	1	**7**	0	0	1	**7**	1	0	0	**7**	1	0	0	**8**	0	0	0	**6**	2	0	0
A24				0	**8**	0	0	0	**8**	0	0	0	**6**	2	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0
A31	2	16	±60	8	0	0	2	**6**	0	1	2	**5**	1	0	**4**	**3**	3	**2**	0	**3**	2	**4**	0	2	2	3	1	**2**
A32				5	2	0	0	**6**	3	1	1	**3**	2	**4**	2	0	0	**8**	0	0	4	**4**	0	0	5	**2**	1	0
A33				3	5	0	0	**3**	**8**	0	0	0	5	1	**2**	0	**7**	1	0	0	**8**	0	0	0	**6**	2	0	0
A34				0	**8**	0	0	0	**7**	1	0	0	**7**	1	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0
A41	2	16	±90	8	0	0	2	**6**	0	0	2	**6**	1	**3**	**2**	**2**	0	**6**	2	0	0	**5**	2	1	1	**4**	2	1
A42				5	4	0	2	**2**	3	1	2	**2**	3	3	**2**	0	2	**6**	0	0	4	**4**	0	0	5	**3**	0	0
A43				3	**6**	0	2	0	**6**	2	0	0	5	1	**2**	0	4	**4**	0	0	**6**	2	0	0	**6**	2	0	0
A44				0	**6**	0	1	1	**7**	1	0	0	**8**	0	0	0	**7**	1	0	0	**8**	0	0	0	**7**	1	0	0
B21	2	36	±30	8	5	0	0	**3**	5	0	0	**3**	5	1	0	**2**	5	**3**	0	0	5	**3**	0	0	**6**	2	0	0
B22				5	**7**	0	0	1	**7**	0	0	1	**6**	1	0	1	5	**3**	0	0	**6**	2	0	0	4	**4**	0	0
B23				3	**7**	0	0	1	**7**	0	0	1	**6**	1	0	1	**7**	1	0	0	**8**	0	0	0	**8**	0	0	0
B24				0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**7**	1	0	0
B31	2	36	±60	8	0	1	1	**6**	2	1	1	**4**	2	**4**	0	2	2	**5**	1	0	1	**6**	1	0	2	**5**	1	0
B32				5	3	1	0	**4**	5	**2**	1	0	3	3	**2**	0	0	**8**	0	0	2	**6**	0	0	4	**4**	0	0
B33				3	**6**	1	0	1	**7**	1	0	0	**6**	2	0	0	5	**3**	0	0	**7**	1	0	0	**7**	1	0	0
B34				0	**7**	1	0	0	**8**	0	0	0	**8**	0	0	0	5	**3**	0	0	**8**	0	0	0	**7**	1	0	0
C11	4	16	±30,	12	4	0	0	**4**	4	0	0	**4**	2	**3**	**3**	0	1	**5**	1	1	0	**5**	2	1	2	**4**	1	1
C12			±60	8	4	0	0	**4**	4	0	1	**3**	0	**3**	**3**	**2**	2	**6**	0	0	4	**4**	0	0	5	**2**	1	0
C13				5	**6**	0	0	2	**6**	0	0	2	**7**	1	0	0	5	**3**	0	0	5	**3**	0	0	5	**2**	1	0
C14				2	**7**	0	0	1	**8**	0	0	0	**6**	2	0	0	**7**	1	0	0	**8**	0	0	0	**8**	0	0	0
C15				0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0
C21	4	16	±30,	8	0	0	2	**6**	1	1	**3**	**3**	0	3	**4**	1	0	**5**	**3**	0	0	**7**	1	0	0	**7**	1	0
C22			±60	7	1	0	2	**5**	2	1	**3**	**2**	1	**5**	2	0	0	**5**	**3**	0	0	**8**	0	0	0	**7**	1	0
C23				6	1	0	2	**5**	4	0	2	**2**	3	**3**	1	1	1	**6**	1	0	4	**4**	0	0	2	**6**	0	0
C24				5	**6**	0	1	1	**7**	1	0	0	**6**	2	0	0	**6**	2	0	0	**7**	1	0	0	**6**	2	0	0
D11	8	16	±15,	13	**7**	0	1	0	**6**	0	2	0	4	1	**3**	0	5	1	**2**	0	**6**	0	2	0	5	0	**3**	0
D12			±30,	10	**7**	1	0	0	**6**	0	2	0	4	1	**3**	0	**6**	1	1	0	**7**	0	1	0	4	2	**2**	0
D13			±60,	7	**8**	0	0	0	**7**	0	0	1	**8**	0	0	0	**7**	1	0	0	**8**	0	0	0	**7**	1	0	0
D14			±90	5	**8**	0	0	0	**8**	0	0	0	**7**	1	0	0	**7**	0	1	0	**7**	1	0	0	**7**	1	0	0
D15				3	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0
D21	8	16	±15,	10	1	0	**7**	0	2	1	**5**	0	0	2	**6**	0	0	**6**	2	0	0	**6**	2	0	0	**4**	**4**	0
D22			±30,	9	3	0	**5**	0	5	0	**3**	0	1	2	**5**	0	1	**6**	1	0	2	**5**	1	0	4	**3**	1	0
D23			±60,	8	**6**	0	2	0	**6**	0	2	0	3	1	**4**	0	3	**5**	0	0	2	**6**	0	0	4	**3**	1	0
D24			±90	7	**8**	0	0	0	**8**	0	0	0	5	**2**	1	0	4	**4**	0	0	5	**3**	0	0	**6**	1	0	1
D25				5	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0	**8**	0	0	0

range of the angle of incidence is spread to ±30°. Generally, L of 5 dB or larger is inclined to cause an image shift. But in the 500 Hz and 1k Hz bands, the most sensitive angles of incidence, i.e., ±90° and ±60°, result in image shift occurrence when L is 3 dB or larger. These sensitive ranges in the angle of incidence in each band correspond well to the most effective angle to decrease the IACC reported by Ando (1985, p. 132). Figure 3 also shows that it is a typical phenomenon of categories perceived moving from 0 to 3 in 125 and 250 Hz bands, while Category 1 lies between Category 0

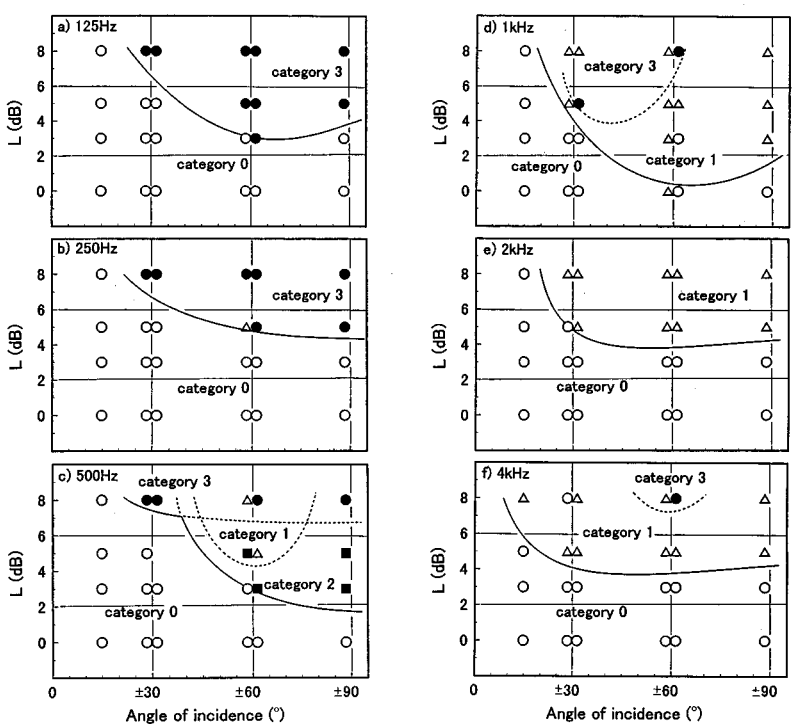


FIG. 3. Representative image shift categories of sound fields composed of the direct sound and two early lateral reflections having symmetrical angles of incidence. The ordinate indicates L and the abscissa indicates the angle of incidence. ○: Category 0, △: Category 1, ■: Category 2, ●: Category 3.

and 3 in 500 Hz and the higher frequency bands. At 500 Hz, Category 2 is more typical than Category 1 as a transient from Category 0 to 3. This subject is described in more detail in the next section. For this kind of complex dependency, some measures are necessary that can evaluate the directional aspect of sound fields in categorization of image shift.

Comparing the cases of A21 to A34 with B21 to B34, it was revealed that the delay in reflections does not cause any major difference in the correlation between the test sound fields and the image shift categories as a whole, except some situations.

B. Dependence on frequency band

Table II and Fig. 3 show that each of the three image shift category representative of each sound field has its specific frequency bands where it is observed, as shown in Table III.

In the 125 and 250 Hz bands, most of the representative image shift categories are Category 3, which is described by the movement of an image from the source direction to the lateral direction. Category 2 sometimes occurs for some cases, especially when both direct sound and reflections with small angles of incidence near the direct sound are faint.

In the 500 Hz band, many of the representative image shift categories are Category 1 or 2. They change into Category 3 for some cases with a small number of reflections when L is about 8 dB or larger. However, Category 2 is observed most frequently.

In the 1k Hz and higher frequency bands, most of the representative image shift categories are Category 1, which is characterized by the appearance of more images in the lateral directions in addition to the source direction. For some cases, especially in the 1k Hz band, when L is about 8 dB or larger, representative images change from Category 1 to 2 or 3.

Many of the test sound fields that correspond to Category 2 or 3 in the 125 Hz and 250 Hz bands correspond to Category 1 in the 1k Hz and higher frequency bands. Few cases are found where the image in the source direction is lost in the higher frequency bands and not lost in the lower frequency bands. This fact indicates that 1k Hz and higher frequency bands are more important in localizing an image to the source direction than 250 Hz and the lower bands when the lateral reflections are very strong. It appears that the 500 Hz band sometimes helps keep an image heard in the source direction although the image disappears in lower bands.

TABLE III. Correspondence between representative image shift categories and frequency bands. ◎: Most of cases fall under the category. ○: Many of cases fall under the category. △: Sometimes occurs when the value of L are about 8 dB or larger. ×: Rarely occurs.

Image shift category	1/1 oct. band center frequency (Hz)					
	125	250	500	1k	2k	4k
1	×	×	△	◎	◎	◎
2	△	△	○	△	×	△
3	◎	◎	△	△	×	×

V. MEASURES DERIVED FROM INTERAURAL CROSS-CORRELATION FUNCTION (IACF) OF TEST SOUND FIELDS

A. Definition of measures derived from IACF

The interaural cross-correlation function (IACF) was measured for each test sound field for the range of delay: $-1\ \text{ms} \leqq \tau \leqq +1\ \text{ms}$, in each of 1/1 octave band with mid-frequencies from 125 to 4k Hz. Three measures for IACF were defined in addition to IACC. Their definitions are shown in Fig. 4.

Peak value of the function for the delay range in each band is IACC as defined by Hidaka *et al.* (1995). IACC was measured for all the test sound fields for each 1/1 octave band from 125 to 4 k Hz.

The first of the three additional measures is τ_0 which is defined by the delay corresponding to the nearest peak of IACF from the central inter-aural delay of $\tau = 0$ ms. It was so defined since, as earlier discussed, the source image position corresponds to the peak position of IACF, as reported in the localization study of Damaske *et al.* (1972) and also in a series of lateralization studies of Bilsen *et al.* (1973), Stern *et al.* (1978, 1988), and Shackleton *et al.* (1991). This measure was derived for all the test sound fields for the 1/1 octave bands of 125, 250, and 500 Hz. For the higher frequency bands, the values were small enough to be identified as 0 ms.

Next, when IACF takes a dip at $\tau = 0$ ms, the case of inter-aural ''opposite phase'' condition, it was denoted as ''opposite phase.'' This occurs when a binaural signal is presented in opposite phase to both ears. Damaske *et al.* (1972) reported that this case produces two separate lateral images at the same time and no image at the front center.

The third measure is $\alpha_{\text{lat.}}$, which is defined by a ratio of two heights of peaks in IACF. The denominator is the height of the nearest peak from the central inter-aural delay of $\tau = 0$ ms. The numerator is the height of the maximum peak for $|\tau| \leqq 1$ ms except the nearest peak from $\tau = 0$ ms. As was described above, peak positions of IACF are important in judging image positions. At 1 k Hz and in the higher bands, IACF also has several peaks within $|\tau| \leqq 1$ ms in addition to the central peak at $\tau = 0$ ms. It corresponds to the possibility of perceiving plural simultaneous images, as reported by Bilsen *et al.* (1973). Considering that normal sound field conditions usually generate a prominent peak at $\tau = 0$ ms, plural simultaneous images are supposed to be perceived when the lateral peaks are as large as the central peak. Furthermore, when the lateral peak is high enough compared to the central peak, only a lateral image will be perceived. $\alpha_{\text{lat.}}$ was derived for the frequency bands of 1k, 2k, and 4k Hz. For the frequency bands of 500 Hz and below, it was not able to be defined because IACF has only one peak for $|\tau| \leqq 1$ ms.

τ_0 and $\alpha_{\text{lat.}}$, are not defined for the cases of ''opposite phase'' because IACF does not have a peak at around $\tau = 0$ ms.

B. Measures of the test sound fields and their correspondence to image shift categories

$[1-\text{IACC}]$, $|\tau_0|$, $\alpha_{\text{lat.}}$, and opposite phase flags for each test sound field in each octave band are shown in Table IV.

In the 125 Hz band, values of $[1-\text{IACC}]$ are in the range from 0 to about 0.4. Values larger than 0.2 are limited in about one-fourth of the whole sound fields. In the 250 Hz band, values of $[1-\text{IACC}]$ are widely distributed from 0 to about 0.9, but only less than one-fourth of the whole cases take the value larger than 0.5. In the 500 Hz band, the opposite phase condition occurs for nearly half the cases of the test sound fields. This condition does not occur in the lower frequency bands at all.

In the 500 Hz and lower bands, $[1-\text{IACC}]$ and $|\tau_0|$ are generally inclined to increase monotonically as L increases.

At 500 Hz among the cases including prominent reflections from $\theta = \pm 60°$ or $\pm 90°$, when $[1-\text{IACC}]$ takes a maximum value of about 0.85 to 0.9, the opposite phase conditions are achieved. Then, $[1-\text{IACC}]$ changes to decrease as L increases keeping opposite phase.

In the 1 k Hz and higher bands, $[1-\text{IACC}]$ also increases as L increases for a range of $\alpha_{\text{lat.}}$ smaller than unity. With the value of L being about 5 dB, $\alpha_{\text{lat.}}$ reaches unity and $[1-\text{IACC}]$ reaches a maximum value of about from 0.85 to 0.9. Then, $[1-\text{IACC}]$ changes to decrease with the increase of L. In several sound fields with L being 8 dB or larger, $\alpha_{\text{lat.}}$ sometimes reaches the range exceeding 2 where the lateral peak of IACF is obviously larger than the central peak. The central peak needs attention to be identified in the cases with the maximum value of 9 in the 1 k Hz band as the peak has a height smaller than 0.1. In several cases, including prominent reflections from $\pm 30°$ in the 1k Hz band, $\pm 15°$ or $\pm 30°$ in the 2k Hz band, $\pm 15°$ in the 4k Hz band, the opposite phase condition is observed with L being about 5–8 dB. Then, $[1-\text{IACC}]$ also changes to decrease with the increase of L.

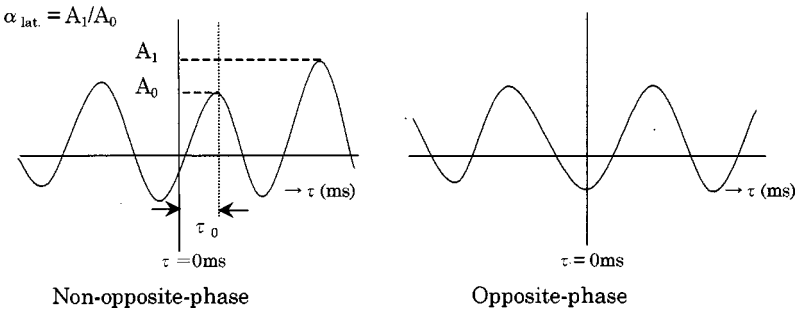


FIG. 4. Definition of measures which are derived from inter-aural cross-correlation function (IACF) of a binaural impulse response of a sound field. Left: Nonopposite phase IACF, usually found in a room's sound field. τ_0 is defined by the delay which correspond to the peak nearest from the central inter-aural delay of $\tau = 0$ ms. $\alpha_{\text{lat.}}$ is defined by the ratio of two heights of peaks: A_0 and A_1, where A_0 is the height of the peak nearest from $\tau = 0$ ms, A_1 is the height of the maximum lateral peak for $|\tau| \leqq 1$ ms except the central peak. Right: Opposite phase IACF, which takes a dip at the central inter-aural delay of $\tau = 0$ ms.

TABLE IV. Measures derived from inter-aural cross-correlation function (IACF) of a binaural impulse responses of test sound fields. N: number of reflections, θ: angles of incidence, Δt_1: initial delay of the first reflection, L: total level of reflections relative to that of the direct sound.

Name of sound field	Attributes of test sound field				1/1 oct. band center frequency																	
					125 Hz			250 Hz			500 Hz			1k Hz			2k Hz			4k Hz		
	N	Δt_1 (ms)	θ (°)	L (dB)	1−IACC	τ_0 (ms)	opp-osite phase	1−IACC	τ_0 (ms)	opp-osite phase	1−IACC	τ_0 (ms)	opp-osite phase	1−IACC	$\alpha_{lat.}$	opp-osite phase	1−IACC	$\alpha_{lat.}$	opp-osite phase	1−IACC	$\alpha_{lat.}$	opp-osite phase
A11	2	16	±15	8	0.03	0.00	0	0.08	−0.03	0	0.27	0.02	0	0.59	0.56	0	0.87	-	1	0.58	-	1
A12				5	0.03	0.00	0	0.07	−0.03	0	0.23	0.02	0	0.52	0.48	0	0.90	1.4	0	0.79	-	1
A13				3	0.02	0.00	0	0.06	−0.03	0	0.19	0.02	0	0.44	0.41	0	0.76	0.82	0	0.91	10	0
A14				0	0.03	−0.05	0	0.06	0.00	0	0.17	0.02	0	0.37	0.41	0	0.61	0.79	0	0.83	0.50	0
A21	2	16	±30	8	0.14	0.05	0	0.26	0.00	0	0.77	0.13	0	0.79	-	1	0.76	-	1	0.67	0.91	0
A22				5	0.11	0.05	0	0.23	0.00	0	0.69	0.07	0	0.90	-	1	0.83	2	0	0.58	0.73	0
A23				3	0.10	0.00	0	0.20	−0.03	0	0.58	0.03	0	0.85	0.73	0	0.76	0.88	0	0.49	0.67	0
A24				0	0.10	−0.10	0	0.17	−0.05	0	0.47	0.03	0	0.69	0.52	0	0.61	0.71	0	0.41	0.62	0
A31	2	16	±60	8	0.22	−0.20	0	0.55	−0.05	0	0.60	-	1	0.73	9	0	0.86	1.67	0	0.80	1.88	0
A32				5	0.19	−0.20	0	0.47	−0.05	0	0.75	-	1	0.76	1.55	0	0.77	0.77	0	0.71	0.55	0
A33				3	0.16	−0.15	0	0.39	−0.05	0	0.87	-	1	0.72	0.80	0	0.62	0.68	0	0.56	0.48	0
A34				0	0.19	−0.10	0	0.36	−0.03	0	0.83	0.12	0	0.59	0.53	0	0.50	0.67	0	0.44	0.48	0
A41	2	16	±90	8	0.46	−0.20	0	0.96	1.00	0	0.42	-	1	0.59	1.53	0	0.70	1.05	0	0.77	0.81	0
A42				5	0.40	−0.15	0	0.86	0.05	0	0.58	-	1	0.61	1.12	0	0.56	0.77	0	0.60	0.43	0
A43				3	0.33	−0.10	0	0.72	0.05	0	0.73	-	1	0.58	0.90	0	0.43	0.75	0	0.48	0.43	0
A44				0	0.27	−0.10	0	0.59	0.00	0	0.89	-	1	0.49	0.72	0	0.33	0.71	0	0.37	0.44	0
B21	2	36	±30	8	0.14	0.05	0	0.26	0.03	0	0.79	0.10	0	0.80	-	1	0.77	-	1	0.67	0.96	0
B22				5	0.12	0.05	0	0.22	0.03	0	0.69	0.07	0	0.90	-	1	0.85	2.5	0	0.58	0.73	0
B23				3	0.10	0.05	0	0.19	0.00	0	0.61	0.03	0	0.83	0.67	0	0.78	0.87	0	0.49	0.67	0
B24				0	0.10	0.00	0	0.16	0.00	0	0.50	0.02	0	0.67	0.50	0	0.62	0.74	0	0.40	0.61	0
B31	2	36	±60	8	0.23	−0.25	0	0.57	−0.07	0	0.58	-	1	0.75	9	0	0.85	1.83	0	0.81	1.75	0
B32				5	0.20	−0.20	0	0.49	−0.05	0	0.75	-	1	0.77	1.33	0	0.74	0.78	0	0.68	0.55	0
B33				3	0.18	−0.15	0	0.41	−0.05	0	0.87	-	1	0.70	0.73	0	0.61	0.71	0	0.54	0.49	0
B34				0	0.17	−0.10	0	0.38	−0.05	0	0.87	0.10	0	0.57	0.52	0	0.48	0.71	0	0.41	0.48	0
C11	4	16	±30,	12	0.16	−0.20	0	0.41	−0.10	0	0.81	-	1	0.86	-	1	0.85	-	1	0.79	1.19	0
C12			±60	8	0.15	−0.15	0	0.37	−0.10	0	0.86	-	1	0.89	-	1	0.90	-	1	0.74	0.88	0
C13				5	0.13	−0.15	0	0.29	−0.07	0	0.84	0.20	0	0.90	1.75	0	0.87	1	0	0.59	0.68	0
C14				2	0.10	−0.10	0	0.21	−0.05	0	0.68	0.06	0	0.82	0.67	0	0.73	0.84	0	0.46	0.61	0
C15				0	0.08	0.00	0	0.13	−0.03	0	0.48	0.03	0	0.69	0.45	0	0.58	0.77	0	0.35	0.58	0
C21	4	16	±30,	8	0.22	−0.40	0	0.50	−0.20	0	0.60	-	1	0.75	6.33	0	0.84	1.7	0	0.78	2.63	0
C22			±60	7	0.21	−0.35	0	0.47	−0.15	0	0.69	-	1	0.77	4	0	0.87	1.5	0	0.81	1.12	0
C23				6	0.19	−0.25	0	0.42	−0.10	0	0.81	-	1	0.81	2.33	0	0.89	1	0	0.75	0.84	0
C24				5	0.14	−0.15	0	0.31	−0.03	0	0.86	0.23	0	0.88	1.5	0	0.88	0.89	0	0.61	0.72	0
D11	8	16	±15,	13	0.12	0.10	0	0.33	0.07	0	0.92	0.35	0	0.85	2.75	0	0.83	-	1	0.89	-	1
D12			±30,	10	0.11	0.10	0	0.30	0.05	0	0.88	0.16	0	0.85	1.57	0	0.89	-	1	0.95	-	1
D13			±60,	7	0.09	0.05	0	0.24	0.03	0	0.75	0.05	0	0.83	0.92	0	0.95	2	0	0.93	1	0
D14			±90	5	0.07	0.05	0	0.17	0.00	0	0.51	−0.06	0	0.71	0.55	0	0.72	1	0	0.66	0.73	0
D15				3	0.05	0.00	0	0.11	0.00	0	0.42	0.03	0	0.68	0.52	0	0.80	0.71	0	0.74	0.54	0
D21	8	16	±15,	10	0.35	−0.20	0	0.75	−0.20	0	0.46	-	1	0.69	2.65	0	0.90	1	0	0.89	1	0
D22			±30,	9	0.28	−0.15	0	0.67	−0.15	0	0.55	-	1	0.71	2.33	0	0.89	1	0	0.88	0.82	0
D23			±60,	8	0.23	−0.10	0	0.62	−0.10	0	0.65	-	1	0.73	1.9	0	0.88	1	0	0.86	0.63	0
D24			±90	7	0.16	−0.05	0	0.49	−0.05	0	0.84	-	1	0.77	1.55	0	0.90	1	0	0.86	0.60	0
D25				5	0.07	0.00	0	0.29	0.03	0	0.76	−0.02	0	0.79	0.87	0	0.91	1	0	0.85	0.64	0

Figure 5 shows plots of representative categories of each test sound field in relation to $|\tau_0|$, opposite phase flag, and [1−IACC]. At 500 Hz, ''opposite phase'' cases are plotted at the right end in abscissa. Open circles, open triangles, filled squares, and filled circles indicate representative categories of 0, 1, 2, and 3, in which figures are underlined in Table II. In the 125 Hz band, sound fields with [1−IACC] equal to or larger than 0.25 are not confined to a specific category. They correspond to various types of categories, both normal and image shift. Sound fields with [1−IACC] smaller than 0.25 are divided into two groups by $|\tau_0|$. Most of the sound fields with $|\tau_0|$ smaller than 0.15 ms correspond to Category 0, and most of the others correspond to Category 3. It is the same in the 250 Hz band, but the boundary values for [1−IACC] and $|\tau_0|$ are 0.5 and 0.1 ms, respectively. In the 500 Hz band, nearly half of the test sound fields are opposite phase. The same kind of pattern as that in the lower two bands is observed when the phase is not opposite. The boundary values for [1−IACC] and $|\tau_0|$ are 0.83 and 0.07 ms, respectively. The opposite phase cases correspond to various types of categories, including Category 2 and the others. The difference in the boundary value for $|\tau_0|$ among bands are considered to correspond to the fact that the lower frequency sound generates larger ASW when IACC and SPL are the same (Setoyama *et al.*, 1995), because larger ASW

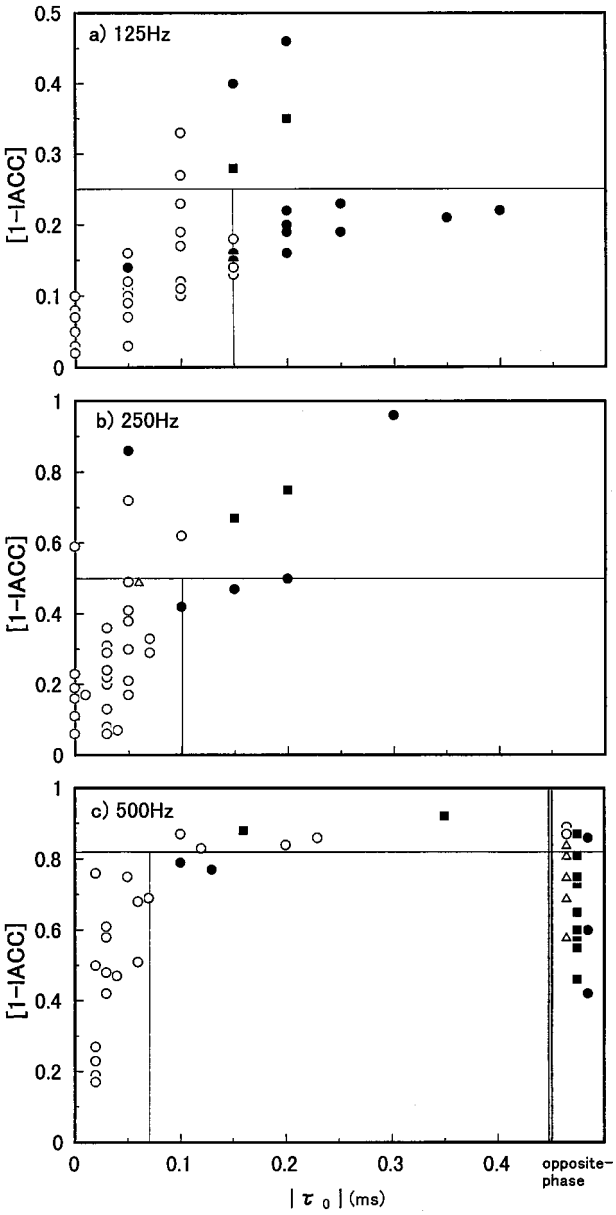


FIG. 5. Plots of representative image shift categories in relation to τ_0, opposite phase flag, and [1−IACC] at the 500 Hz band and in the lower bands. ○: Category 0; △: Category 1; ■: Category 2; ●: Category 3.

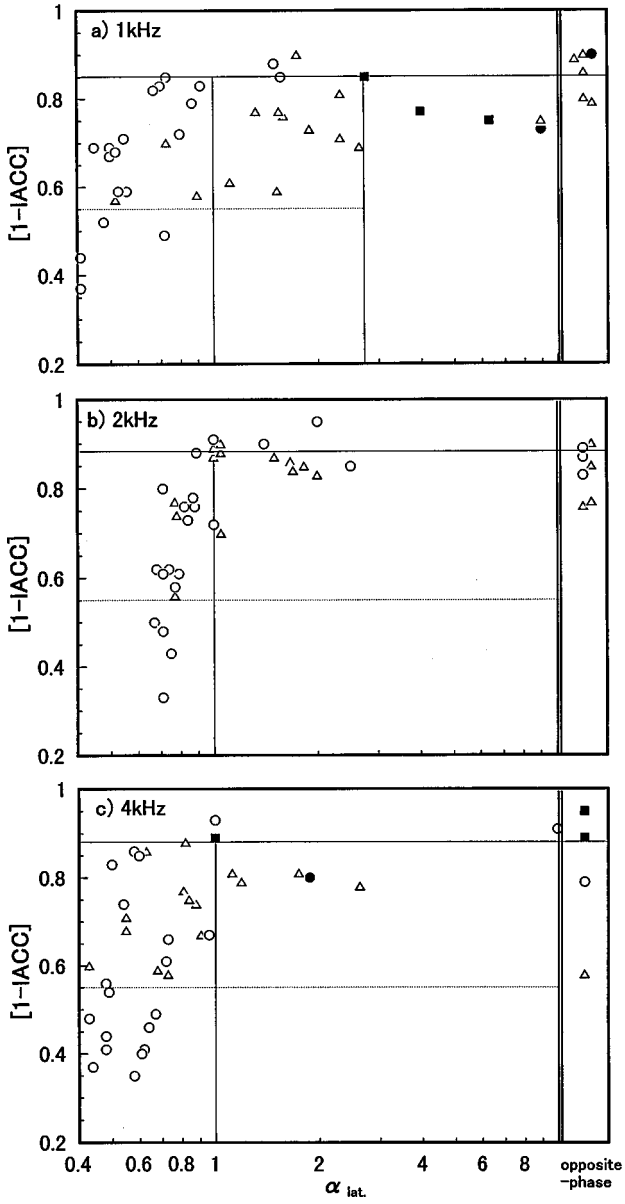


FIG. 6. Plots of representative image shift categories in relation to $\alpha_{lat.}$, opposite phase flag, and [1−IACC] in the 1k Hz and the higher bands. ○: Category 0; △: Category 1; ■: Category 2; ●: Category 3.

needs a larger shift in direction in order not to include the source direction.

Figure 6 shows the same kind of plots in relation to $\alpha_{lat.}$, opposite phase flag, and [1−IACC] in 1k Hz and above bands, in the same manner as Fig. 5. Also, the opposite phase cases are plotted at the right end in abscissa. In these bands, only a part of the test sound fields are opposite phase in contrast to the 500 Hz band. Corresponding categories are not limited to a specific category when [1−IACC] is larger than a certain value, which is 0.85 in the 1k Hz band and 0.88 in the 2k Hz and 4k Hz bands. In the 1k Hz band, the other area is divided into three with $\alpha_{lat.}$ smaller than unity, between unity and 2.7, and larger than 2.7, respectively. Most of cases with $\alpha_{lat.}$ smaller than unity fall under Category 0. Within this area, all of the plots with [1−IACC] smaller than 0.55 are of Category 0. All of the cases with $\alpha_{lat.}$ between unity and 2.7 fall under Category 1. The area of $\alpha_{lat.}$ larger than 2.7 mostly fall under Categories 2 and 3, where the possibility of an image in the source direction disappearing is considerably high. Those sound fields tend to have the image disappear in the source direction not only in the 1k Hz band but also in the 500 Hz and lower frequency bands. The interpretation of plots is almost the same for 2k Hz and 4k Hz bands. However, no plots were found where $\alpha_{lat.}$ is larger than 2.7. For these three bands, the opposite phase cases correspond to Category 0 or 1 when [1−IACC] is smaller than 0.85 in the 1k Hz band and 0.88 in the 2k and 4k Hz bands.

From the above observation, correspondence between measures and categories of source image has been identified as follows. $|\tau_0|$ is a measure for the 500 Hz and lower bands, which corresponds to the shifting from Category 0 to 3. $\alpha_{lat.}$ is a measure for the 1k Hz and higher frequency bands, which corresponds to the shifting from Category 0 to 1, and

TABLE V. Discrimination of representative image shift categories by measures of image shift. Underline figures indicate those representative image shift categories that are most frequently observed in each condition of an additional measure. Hit zones of category discrimination between [0,1] and [2,3] by $[1-\mathrm{IACC}_E]$ and the additional measures are indicated by the bolded figures.

Center frequency	Measures of image shift and the criteria		Number of sound fields corresponding to each representative categories				Hit rate of discriminating categories [0,1] and [2,3]
	$[1-\mathrm{IACC}_E]$	additional	0	1	2	3	
125 Hz	$\geqq 0.25$		2	0	**2**	**2**	
	<0.25	$\tau_0 \geqq 0.15$ ms	3	0	**0**	**10**	0.84
		$\tau_0 < 0.15$ ms	**22**	**0**	0	2	
250 Hz	$\geqq 0.5$		3	0	**2**	**5**	
	<0.5	$\tau_0 \geqq 0.10$ ms	0	0	**0**	**4**	0.86
		$\tau_0 < 0.10$ ms	**25**	**1**	0	3	
500 Hz	$\geqq 0.83$		5	2	**3**	**1**	0.71
	<0.83	$\tau_0 \geqq 0.07$ ms	1	0	**0**	**2**	
		opposite phase	0	4	**7**	**3**	
		$\tau_0 < 0.07$ ms	**14**	**0**	0	0	
1k Hz	$\geqq 0.85$		3	4	**1**	**1**	0.81
	<0.85	$\alpha_{\mathrm{lat.}} \geqq 2.7$	0	1	**2**	**1**	
		opposite phase	**0**	**2**	0	0	
		$1 < \alpha_{\mathrm{lat.}} < 2.7$	**0**	**9**	0	0	
		$\alpha_{\mathrm{lat.}} < 1$	**16**	**3**	0	0	
2k Hz	$\geqq 0.88$		5	6	**0**	**0**	0.74
	<0.88	opposite phase	**2**	**3**	0	0	
		$\alpha_{\mathrm{lat.}} > 1$	**2**	**7**	0	0	
		$\alpha_{\mathrm{lat.}} < 1$	**15**	**3**	0	0	
4k Hz	$\geqq 0.88$		2	1	**3**	**0**	0.91
	<0.88	opposite phase	**1**	**1**	0	0	
		$\alpha_{\mathrm{lat.}} > 1$	**0**	**4**	0	1	
		$\alpha_{\mathrm{lat.}} < 1$	**19**	**11**	0	0	

further into Categories 2 or 3 in some extreme cases, provided that Category 0 and 1 are not always divided clearly in some cases, especially in the 4k Hz band. The opposite phase flag is a measure primary for the 500 Hz band, which corresponds primary to the shifting from Category 0 to 2 as well as to 1 and 3 in some cases. This phenomenon is observed most frequently when the angles of incidence include $\pm 90°$. For those angles, lateral reflections reach both ears in the opposite phase at 500 Hz because the path difference between ears half round the head coincides approximately with half of the wavelength. When $[1-\mathrm{IACC}]$ is larger than a certain value defined differently for each frequency band and 0.85 averaged for the 500 Hz and higher bands, the category is not predictable by the measures above. This condition fall under Categories 2 and 3 at a considerable frequency. These descriptions above are summarized in Table V. This table also shows the effectiveness of the measures in discriminating two larger categories [0,1] and [2,3]. Category [0,1] produces an image to the source direction, while Category [2,3] does not. The consistency of the larger categories with practical judgement will be discussed in the next section. The hit rate of the discrimination is in the range from 0.71 to 0.91, which is averaged as 0.81 over the six frequency bands. Effectiveness of the measures is thus confirmed.

VI. MEASURES OF IMAGE SHIFT IN EXISTING CONCERT HALLS

In order to determine practical criteria in image shift judgement, $[1-\mathrm{IACC}_E]$, $|\tau_0|$, $\alpha_{1\mathrm{at.}}$ and opposite phase flags were derived from the early 80 ms part of binaural impulse responses measured with an omni-directional source in three existing concert halls. They were selected from properly designed halls to cover the range from middle to highest class concert halls often used by the best orchestras of the world. One of the three halls is ranked as $[\mathrm{A}^+$, ''Superior''] and the other two as $[\mathrm{B}^+$, ''Good to Excellent''] by Beranek (1996). These indicate the acceptable range for the measures. Two of them are shoebox shaped and the third not. All of them have more than 2000 seats. Average room volume for all three is approximately 2000 m^3. Early 80 ms of impulse responses, including the direct sound and the portion after its arrival, were used to derive IACF. As mentioned earlier, the early 80 ms part is a major contributor to creation of apparent source images and is important in evaluating concert hall acoustics (Hidaka *et al.*, 1995). IACC_E indicates that the value is derived from the early 80 ms part of impulse responses.

Figure 7 shows the results for 500 Hz and lower frequency bands. Closed circles show plots of $[\mathrm{A}^+]$ ranking shoebox type halls, while open squares and open triangles show plots of $[\mathrm{B}^+]$ ranking halls. Among the two $[\mathrm{B}^+]$ halls,

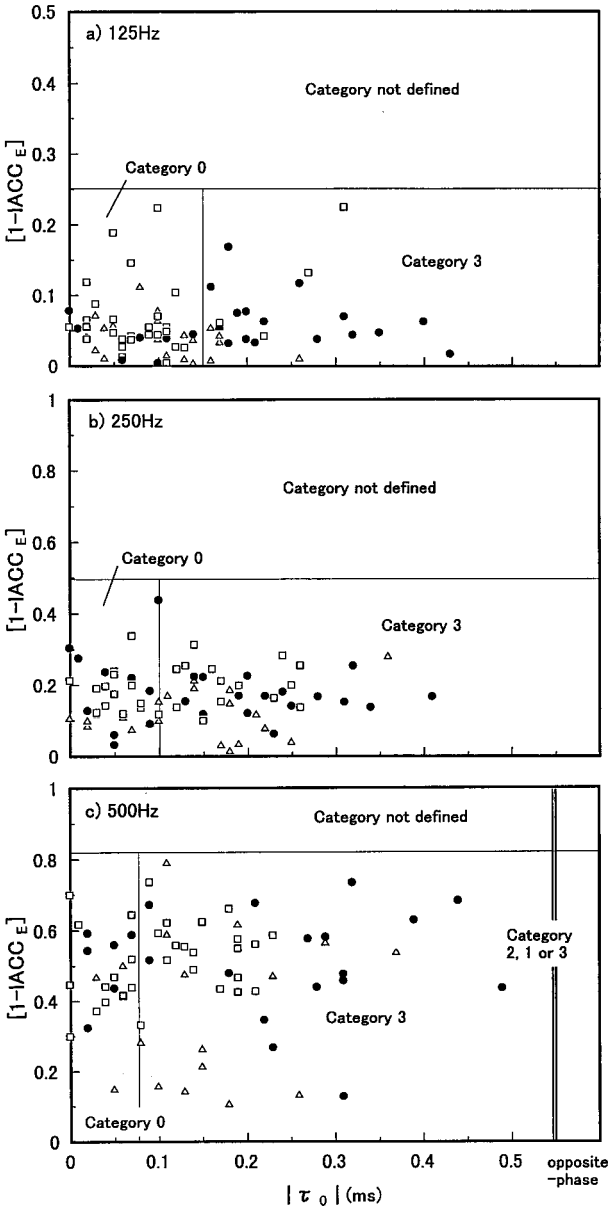


FIG. 7. Plots of τ_0, opposite phase flag, and $[1-\mathrm{IACC}_E]$ measured in three existing concert halls at 500 Hz and the lower bands. ●: $[A^+]$ ranked shoebox hall; □: $[B^+]$ ranked shoebox hall; △: $[B^+]$ ranked nonshoebox hall.

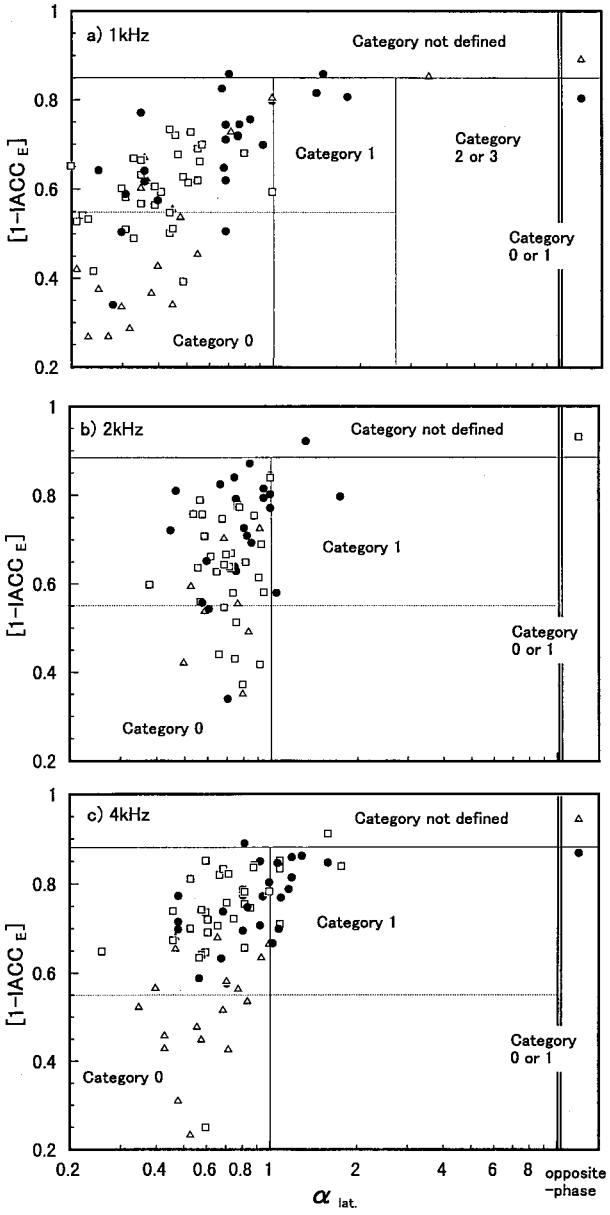


FIG. 8. Plots of α_0, opposite phase flag, and $[1-\mathrm{IACC}_E]$ measured in three existing concert halls in 1k Hz and the higher bands. ●: $[A^+]$ ranked shoebox hall; □: $[B^+]$ ranked shoebox hall; △: $[B^+]$ ranked nonshoebox hall.

open squares show plots of a shoebox type hall. Positions of sources and receivers are shown in the Appendix of Hidaka *et al.* (1995). The areas found to fall under specific categories in the previous section are also indicated in the same manner as Fig. 5. The plots are spread over wide areas corresponding to Category 0 and 3 regardless of the ranking or the room shape. This fact suggests that a directional deviation of a source image from the source direction in these lower frequency bands, corresponding to a large value of $|\tau_0|$, does not always result in reduction of acoustical quality, assuming that there are high frequency components in music that can produce the source image in the right direction. No plots were found opposite phase at 500 Hz among these $[A^+]$ and $[B^+]$ halls.

In Fig. 8, for the 1k Hz band and above, most of plots are placed within the area where $\alpha_{\mathrm{lat.}}$ is equal to and smaller than unity. Many of plots for nonshoebox halls have smaller $[1-\mathrm{IACC}_E]$ than 0.55 in each of 1k, 2k, and 4k Hz bands. As confirmed by the prior studies, the $[A^+]$ ranked hall has the largest average $[1-\mathrm{IACC}_E]$. Compared with the two shoebox type halls, the $[A^+]$ hall has larger $\alpha_{\mathrm{lat.}}$, especially in the 1k Hz band. The differences between these three halls are the most prominent in the 1k and 2k Hz bands. Opposite phase cases are limited to a few seat positions in both $[A^+]$ and $[B^+]$ ranking halls. Some plots for the $[A^+]$ hall are in the range of $\alpha_{\mathrm{lat.}}$ between unity and 2.7. The excess of unity in $\alpha_{\mathrm{lat.}}$ for $[A^+]$ halls is principally observed at two typical seat positions. One is the seat on the rear parquet, and the other a side balcony seat near the stage. At these seat positions, the number of lateral reflections is maximized at the rear parquet seat and the direct sound is inclined to be dif-

TABLE VI. Acceptable range for measures to avoid ''image shift.'' The principal conclusion is that neither $[1-\mathrm{IACC}_E]$ nor $\alpha_{\mathrm{lat.}}$ should exceed the values shown and that opposite-phase should not occur in the band shown, when measured at individual seats.

Measure	1/1 oct. band center frequency (Hz)			
	500	1k	2k	4k
$1-\mathrm{IACC}_E$	<0.83	<0.85	<0.88	<0.88
$\alpha_{\mathrm{lat.}}$	-	<2.75	-	-
opposite phase	not ''opposite phase''	-	-	-

fracted by the balcony front for the side balcony seat. These conditions are considered to make $\alpha_{\mathrm{lat.}}$ larger in those seat positions. One plot for the $[\mathrm{B}^+]$ nonshoebox hall has larger $\alpha_{\mathrm{lat.}}$ than 2.7 at 1k Hz. Although this also occurs in the side balcony seat, it is also considered to be the reason that the hall does not have sufficient fine diffusion on the walls and the ceilings compared with the other two.

$\alpha_{\mathrm{lat.}}$ between unity and 2.7, observed in some seats in the $[\mathrm{A}^+]$ hall, falls under a condition where separate sound images are observed in addition to the source direction but the image in the source direction does not yet disappear. The same kind of cases would be much frequently observed in smaller chamber music halls. It would be necessary to consider the directivity of an instrument to determine what kind of image the audience sitting in such a seat really would perceive, but an omni-directional source is considered as to what represents various types of musical instruments of varying directivity placed on the stage and the variety of their arrangement on the stage. Rather, source images of other frequency bands and that by the subsequent reverberation might cover some gap between the center and the lateral images in those cases. It would therefore be concluded that $\alpha_{\mathrm{lat.}}$ between unity and 2.7 in some seats, which corresponds to Category 1, is acceptable for a highly evaluated concert hall. $\alpha_{\mathrm{lat.}}$ exceeding 2.7, which is observed in the $[\mathrm{B}^+]$ ranking nonshoebox hall, may not be judged as an excellent attribute as the range fall under categories which do not always produce the image in the source direction, i.e., Category 2 and 3.

Summarizing the above discussion in relation to the findings in the previous section, the acceptable range for a concert hall in which no image shift will occur is suggested in Table VI. The upper limit for the preferable range of $[1-\mathrm{IACC}_{E3}]$ is calculated as 0.85 by averaging the critical values for $[1-\mathrm{IACC}_E]$ in three bands of 500, 1k, and 2k Hz.

VII. CONCLUSION

$[1-\mathrm{IACC}]$ and three additional measures, $\alpha_{\mathrm{lat.}}$, τ_0, and the opposite phase flag, which cover the bands ranging from 125 to 4k Hz in relation to three image shift categories, are found to correlate to the occurrence of ''image shift.''

Five hundred Hz and higher frequency bands, especially the 1k Hz band and the higher frequency bands, are very important in keeping the image heard in the source direction when the lateral reflections are very strong. Two-hundred and fifty Hz and the lower frequency bands are not so effective in evaluating concert halls in relation to the image shift.

$\alpha_{\mathrm{lat.}}$ smaller than 2.7 at 1k Hz, not ''opposite phase'' at 500 Hz, and $[1-\mathrm{IACC}_E]$ smaller than 0.83 in the 500 Hz band, 0.85 in the 1k Hz band, 0.88 in the 2k Hz and 4k Hz bands are suggested as a set of practical conditions, under which one can avoid occurrence of the image shift. It allows objective judgement on the image shift to be made in the tuning period of a hall before the opening or during the design stage by measurements in a precisely built scale model.

It is concluded that when $[1-\mathrm{IACC}_{E3}]$, which is one of measures for evaluation of a concert hall in relation to ASW, is 0.85 or larger in some individual seats, as the average value over three bands of 500, 1k, and 2k Hz, that hall should not be judged as always having excellent conditions in terms of acoustics and the acoustic ranking.

ACKNOWLEDGMENTS

The author thanks Dr. Yoichi Ando, professor of Kobe University, for the valuable discussion the author had with him during this research work. The author also thanks Dr. Leo L. Beranek for his suggestions in completing this research.

Ando, Y. (**1977**). ''Subjective preference in relation to objective parameters of music sound fields with a single echo,'' J. Acoust. Soc. Am. **62**, 1436–1441.

Ando, Y. (**1985**). *Concert Hall Acoustics* (Springer-Verlag, Berlin).

Barron, M. (**1971**). ''The subjective effects of first reflections in concert halls—The need for lateral reflections,'' J. Sound Vib. **15**, 475–494.

Barron, M. (**1993**). *Auditorium Acoustics and Architectural Design* (E & FN Spon, London).

Beranek, L. L. (**1996**). *How They Sound: Concert and Opera Halls* (Acoustical Society of America, Woodbury, NY).

Bilsen, F. A., and Raatgever, J. (**1973**). ''Spectral dominance in binaural lateralization,'' Acustica **28**, 131–132.

Damaske, P., and Ando, Y. (**1972**). ''Interaural crosscorrelation for multichannel loudspeaker reproduction,'' Acustica **27**, 232–238.

Hidaka, T., Beranek, L. L., and Okano, T. (**1995**). ''Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls,'' J. Acoust. Soc. Am. **98**, 988–1007.

Okano, T., Beranek, L. L., and Hidaka, T. (**1998**). ''Relations among interaural cross-correlation coefficient (IACC_E), lateral fraction (LF_E), and apparent source width (ASW) in concert halls,'' J. Acoust. Soc. Am. **104**, 255–265.

Setoyama, H., Sakagami, K., Morimoto, M., and Iida, K. (**1995**). ''Frequency dependence of auditory source width,'' Proceedings of the Spring Meeting of Acoustical Society of Japan (in Japanese).

Shackleton, T. M., Medddis, R., and Hewitt, M. J. (**1992**). ''Across frequency integration in a model of lateralization,'' J. Acoust. Soc. Am. **91**, 2276–2279.

Stern, R. M., and Colburn, H. S. (**1978**). ''Theory of binaural interaction based on auditory-nerve data. IV. A model for subjective lateral position,'' J. Acoust. Soc. Am. **64**, 127–140.

Stern, R. M., Zeiberg, A. S., and Trahiotis, C. (**1988**). ''Lateralization of complex binaural stimuli: A weighted-image model,'' J. Acoust. Soc. Am. **84**, 156–165.

Yost, W. A., and Gourevitch, G. (**1987**). *Directional Hearing* (Springer-Verlag, New York).

Method to resolve microphone and sample location errors in the two-microphone duct measurement method

Brian F. G. Katz[a)]
Acoustics Department, The Pennsylvania State University, State College, Pennsylvania 16801

(Received 5 January 1999; revised 18 September 1999; revised 17 May 2000; accepted 7 August 2000)

Utilizing the two-microphone impedance tube method, the normal incidence acoustic absorption and acoustic impedance can be measured for a given sample. This method relies on the measured transfer function between two microphones, and the knowledge of their precise location relative to each other and the sample material. In this article, a method is proposed to accurately determine these locations. A third sensor is added at the end of the tube to simplify the measurement. First, a justification and investigation of the method is presented. Second, reference terminations are measured to evaluate the accuracy of the apparatus. Finally, comparisons are made between the new method and current methods for determining these distances and the variations are discussed. From this, conclusions are drawn with regards to the applicability and need for the new method and under which circumstances it is applicable. Results show that the method provides a reliable determination of both microphone locations, which is not possible using the current techniques. Errors due to inaccurate determination of these parameters between methods were on the order of 3% for $|R|$ and 12% for $\mathrm{Re}\{Z\}$. © *2000 Acoustical Society of America.* [S0001-4966(00)02311-0]

PACS numbers: 43.58.Bh, 43.55.Ev, 43.20.Fn, 43.60.Qv [SLE]

I. INTRODUCTION

The measurement of acoustic absorption and acoustic impedance has been used for decades in the understanding of acoustics and material behaviors. The classical method using pure tones and a standing wave tube has, until recently, been the only reliable method for performing these measurements. The primary limitation of this classical method is the large amount of time needed for each frequency measurement. In the last few decades, measurement technology has advanced and now the two-microphone transfer function method for measuring acoustic absorption and impedance has become much more common. There have been a number of papers describing this method and examples of results.[1–6]

The two-microphone method, implemented here using random noise excitation, has a number of limitations, some similar to the standing wave ratio method, and some different. In both schemes, the distance from the sample to the microphone, either fixed or mobile as in the standing wave tube, is an essential component in the calculation of the acoustic properties of the test material. Bodén and Åbom examined the effects of some types of errors in the two-microphone technique.[1] They discussed the potential percentage variation in the results due to errors in measuring these distances. One method is suggested for determining the distance between the microphones in an attempt to reduce these errors. No suggestion is made for determining the distance to the sample other than using a ruler.

The measurement standard requires knowledge of these locations to an accuracy of 0.01 cm.[7] The purpose of this article is to examine this requirement, its effect, and its attainment. A new method for determining the location of both microphones to the accuracy required by the measurement standard is presented. A third sensor is used at the termination position to assist in resolving the location of the measurement microphones to the necessary precision. The effect of errors in the locations of the sensors is investigated for calculations of the reflection coefficient and impedance.

II. THE STANDARD TWO-MICROPHONE APPROACH

A. Theory

A brief overview of the two-microphone method is presented here. This method utilizes the measured transfer function, h_{12}, between two microphones to separate the incident (p_i) and reflected (p_r) pressure waves. From this, it is possible to calculate the complex reflection coefficient, R, at the measured surface. The complex impedance, Z, can be derived from this calculation. A plane wave tube was utilized in this research, configured as shown in Fig. 1. The theory behind this measurement can be described in the following manner.

The measured transfer function in the frequency domain, H_{12}, is defined by

$$H_{12}=\frac{P_2}{P_1}=\frac{P_{2i}+P_{2r}}{P_{1i}+P_{1r}}, \tag{1}$$

where P is the Fourier transform of $p(t)$, and P_{ni} and P_{nr} are the pressures due to the incident and reflected waves at sensor n. It is then possible to define the following:

$$H_{12i}=\frac{P_{2i}}{P_{1i}}, \quad H_{12r}=\frac{P_{2r}}{P_{1r}}, \quad R_1=\frac{P_{1r}}{P_{1i}}, \quad R_2=\frac{P_{2r}}{P_{2i}}. \tag{2}$$

[a)]Currently with Arup Acoustics, 155 Avenue of the Americas, New York, NY 10013.

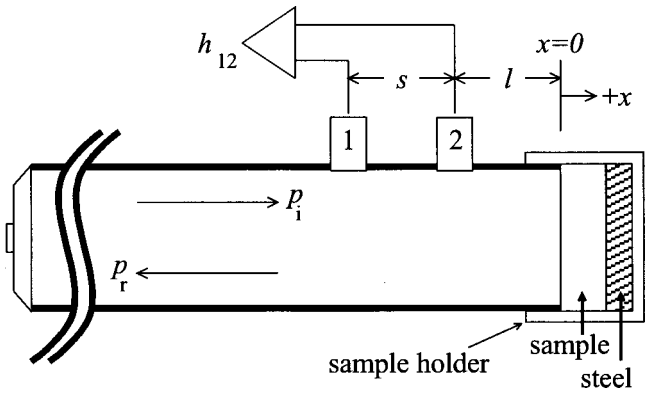


FIG. 1. Impedance tube measurement setup showing measurement microphones $Ch1$ and $Ch2$. Define s as the spacing between the sensors and l as the distance from the second sensor to the impedance surface.

After a little algebra the reflection coefficient at microphone position 1, $Ch1$, can be shown to be

$$R_1 = \left[\frac{H_{12} - H_{12i}}{H_{12r} - H_{12}}\right]. \tag{3}$$

If measurements are limited to frequencies below the first cross mode frequency, only plane waves propagate, and the transfer functions H_{12i} and H_{12r} are simply e^{-jks} and e^{+jks}, respectively. In addition, translation of the calculated reflection coefficient to the surface $x=0$ is only a multiplier of $e^{j2k(l+s)}$. When combined the result is

$$R = e^{j2k(l+s)} R_1 = -e^{j2k(l+s)}\left[\frac{H_{12} - e^{-jks}}{H_{12} - e^{+jks}}\right]. \tag{4}$$

The absorption coefficient is calculated simply as $\alpha = 1 - |R|^2$. The surface impedance of the material is finally determined using Eq. (5) where Z_s is the specific acoustic impedance of the surface and Z is the corresponding specific acoustic impedance ratio.

$$Z = \frac{Z_s}{\rho_0 c} = \frac{1+R}{1-R}. \tag{5}$$

B. Apparatus

The parameters of the system, tube diameter, and sensor locations were designed to obtain optimum results over the frequency range of interest. The chosen frequency range of interest was determined to be 1–6 kHz. The center frequency of this range was 3.5 kHz. The justification for this range is related to future work on the system, but can be summarized as the need to measure acoustic absorption and impedance in the frequency region of speech. It was also decided that the sensor positions would be fixed, allowing for only a single microphone and sample spacing.

The measurement method requires that only plane wave propagation occurs in the tube. For this reason, the upper frequency range is limited by the first cross mode of the tube. The ASTM standard[7] requires the diameter to be less than $0.586c/f_{\max}$ which, for an upper limit of 6 kHz, gives a maximum tube diameter of 3.35 cm. A 1-in. PVC tube was chosen. The internal diameter of the tube was 2.60 cm, with a wall thickness of 0.35 cm. A tube length of 1 m is long enough to ensure that only plane wave propagation is present. It is desirable to have the microphones as close as possible to the termination to obtain the greatest signal-to-noise ratio. The distance to the termination is limited by the presence of evanescent waves coming from the impedance termination, which is not necessarily perfectly planar. The standard suggests placing the near sensor no closer than one tube diameter away. A distance of two internal tube diameters was chosen here for the distance between microphone position 2, $Ch2$, and the termination, resulting in a desired distance of $l = 5.2$ cm. The sensor separation distance is defined in the standard such that $s \ll c/(2f_{\max})$ resulting in a maximum spacing of 2.86 cm. The optimal spacing, resulting in the minimal amount of error, is defined by choosing $s = c/(4f^2_{\text{center}})$.[1] Using 3.5 kHz as the center frequency of interest results in a desired separation distance of $s = 2.45$ cm. The dimensions of the tube correspond to a value of $ka = 1$ at 2.1 kHz.

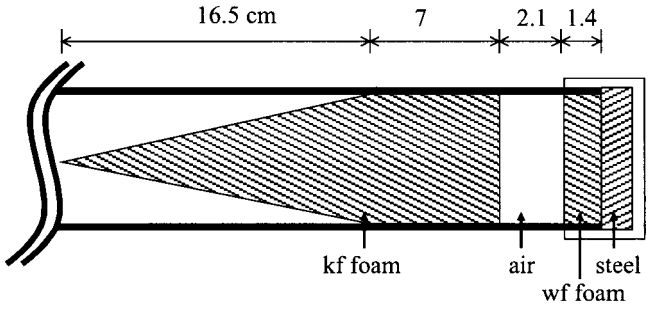


FIG. 2. Anechoic termination construction showing conical wedge.

The system utilized miniature Sennheiser microphone capsules (model KE 4-211-2), a 2.5 cm (1 in.) diameter speaker, electronic preamplifiers, and a multi-channel Hewlett Packard spectrum analyzer. The sensor microphones used have an exposed diaphragm diameter of approximately 0.1 cm. The sensors are flush mounted in the tube, using a clay seal on the outside to ensure against leakage. An anechoic termination was constructed, as shown in Fig. 2.

C. Microphone calibration

The two microphones used in the impedance tube were calibrated relative to each other using the standard switching technique. For this setup, the transfer function was measured for the microphones in the measurement position, then the sensor positions were switched and the transfer function was measured. The geometric mean of the transfer functions is used as the calibration function in all subsequent calculations, where $H_{\text{cal}} = \sqrt{H_{12} \times H_{21}}$ and $H_{12\text{calibrated}} = H_{12\text{measured}} \div H_{\text{cal}}$. This ensured that any magnitude and phase variations in the measurement were a function of the sound field, and not due to differences between the sensors. This procedure was performed prior to each measurement session, resulting in a separate calibration function for each data set to account for any possible effects due to ambient room conditions or drift of sensor characteristics.

III. RESOLUTION OF APPARATUS PARAMETERS

A. Microphone distance to termination

According to the ASTM standard for the two-microphone impedance tube measurement, the distance from the end microphone to the surface sample must be known to an accuracy of 0.01 cm.[7] In the standard this requirement is waived for highly absorptive, roughly textured surfaces. If we chose to measure materials that are not highly absorptive, then this requirement must be fulfilled.

It is therefore necessary to determine the distance from this microphone to the rigid or semirigid termination. The diaphragm, however, had an exposed diameter on the order of 0.1 cm, which is much larger than the prescribed location accuracy of 0.01 cm. Therefore, even if it were possible to measure the physical distance with a caliper or laser scheme, the precision required is on a scale smaller than the size of the diaphragm. It is therefore necessary to determine the distance from the *acoustic* center of the closer diaphragm to the termination.

A standard way to determine this position, from the standing wave ratio tube method, is to move the position of the microphone, and scan for pressure minima. This method contains possible errors in determining the exact acoustic position of the end of the tube, requiring end corrections to be determined and included. From this position, it is possible to calculate this distance quite easily. In designing this system, the position of the microphones was determined, and then fixed. In addition, due to the small tube diameter, the insertion of a probe microphone inside the tube would most likely disturb the sound field in the tube such that the plane wave assumption would be compromised. Therefore, a method which allows for the fixed position of the microphone must be devised. Similar to the traditional method of scanning the sound field for the pressure minima at a given frequency and transforming the problem to the frequency domain, it should be possible to perform the same measurement by sweeping frequency and determining the exact frequency at which the fixed microphone is at a pressure minima. The necessary frequency precision was determined as follows: the approximate distance to the termination was measured to be 5 cm. Using 343 m/s as an approximate sound speed, this would result in a first null (for a quarter wavelength) to be at 1715 Hz. A variation of 0.01 cm would yield a distance of 5.01 cm, and a quarter wavelength frequency of 1711.6 Hz. From this rough calculation, a frequency resolution of 3.4 Hz is necessary.

The accuracy of the acoustic method for determining the position of the microphones is limited, in part, by an accurate determination of the sound speed. The speed of sound was calculated taking into account measures of the temperature, barometric pressure, and relative humidity and using a polynomial approximation of published data and atmospheric effects on sound speed as described in Pierce.[8] Using a rough calculation based on a frequency of 1.7 kHz with a wavelength of 20.18 cm, the resulting changes in conditions to achieve an error of 0.01 cm are $\Delta T = 2.7\,^{\circ}\mathrm{C}$, $\Delta P = 16.9\,\mathrm{mbar}$, or $\Delta RH = 85\%$. These are all well within the measurement accuracy of standard laboratory equipment.

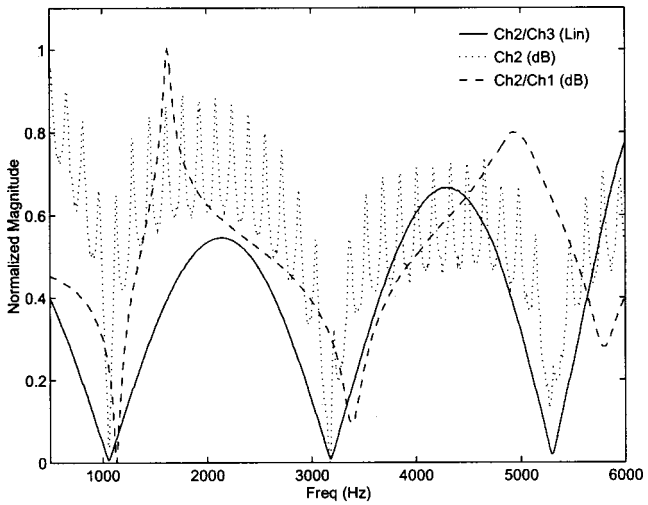


FIG. 3. Potential data for distance calculation: raw impedance sensor sound pressure ($Ch2$); transfer function between one impedance microphone, ($Ch2$) and the termination microphone ($Ch3$); and transfer function between the two impedance microphones ($Ch1$ and $Ch2$). The *y*-scale has been normalized for both the linear and log data so that all the data sets can be visually compared.

Sweeping the frequency at a resolution of 3.4 Hz is possible, but, due to the fact that the exact null is hidden within the noise floor (which results in a minima that spans approximately a 10–15-Hz range), precise determination of this frequency would not be possible by simply choosing the frequency at which the null occurs. Instead, interpolation of the data curve surrounding the null was used to extrapolate the frequency null and avoid contamination of the prediction by the noise floor. In doing this, it was also possible to sweep frequencies at a resolution of 4 Hz as the curve shape, not the exact frequency of any one specific data point, is important.

For confidence in the interpolation, the shape of the frequency spectrum data should be smooth in the region before and after the null. The raw pressure data, measured as a function of frequency for either of the two measurement microphones, contains a dense resonant structure of the tube system. These resonances are in addition to the nulls due to the distance from the termination. An example of a raw microphone measurement is shown by the dotted curve in Fig. 3. It is clear to see that exact determination of the null position would be difficult using the raw pressure data, due to the influence of the tube resonances. These resonances depend on termination impedance, tube length, and driver, but not the sensor location. In order to remove this corruptive influence so that the measurement contains only distance-dependent information useful for determining the distance to the termination, a microphone was placed flush with the termination. A termination consisting of a steel disc with a microphone (designated $Ch3$) flush mounted in the center was constructed for this measurement, as shown in Fig. 4. The termination sensor was sealed in position and, since it was located at the termination, contained the resonance information of the system, without any interference from location specific modal effects. The change in the termination impedance due to the presence of the miniature sensor is negligible to the distance calculations. The effect would be a

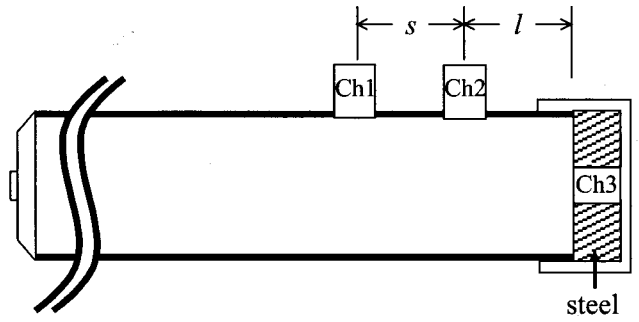


FIG. 4. Modified impedance tube setup showing the two standard measurement microphones, $Ch1$ and $Ch2$, and the additional third termination sensor, $Ch3$, used for determining the measurement microphone locations.

very minor shift in the location of the frequency nulls, less than the 3.4-Hz resolution needed.

A measured transfer function between one of the measurement microphones and the termination microphone yields the necessary data (see $Ch2/Ch3$ curve in Fig. 3). Several nulls exist over the selected frequency range, each one capable of being used to determine the distance. This allows for error checking and redundancy in the distance calculation. Use of the transfer function between the impedance sensors, H_{12} (curve $Ch2/Ch1$ in Fig. 3), containing distinct peaks and nulls, is not suitable for determining distances. As seen in Fig. 3, the peak and null positions for H_{12} are shifted up in frequency from the correct positions and would not be useful in distance determination. This shift can be seen in relation to both the raw pressure measurement ($Ch2$) and the transfer function with respect to the termination sensor ($Ch2/Ch3$).

Interpolation of the frequency minima was determined using the termination sensor transfer functions. A second-order polynomial fit was utilized for the left and right side of each minima, where the intersection produced the result. From this position (in frequency) the distance was calculated depending on the null ($\frac{1}{4}$, $\frac{3}{4}$, or $\frac{5}{4}$ wavelength null) and as a function of the speed of sound, determined from the air properties in the experimental chamber at the time of measurement.

The interpolation method used to determine the frequency null minima was repeatable for a given frequency null within approximately 1.5 Hz. Details of the interpolation method can be found in the thesis by Katz.[9] Several measurement trials were performed with the results given in Table I with the corresponding null frequencies. Taking an average for the distances determined from each $Ch2$ null from Table I, the distance l was determined to be 5.31 cm.

B. Spacing between microphones

Using the distance to termination data for both measurement microphones it is possible to determine the spacing between the microphones, s, a distance necessary for the acoustical calculations. Using the results from Table I and taking an average for the distances determined from each null, the separation distance was determined to be 2.37 cm. This distance can also be calculated by using an anechoic termination and determining the phase delay between the two sensors, again using the transfer function.[1] This method of measurement is totally independent of the previous method and was therefore used as a reference check.

TABLE I. Microphone distance to termination calculations for each null in measurement ($Ch1$:3 nulls, $Ch2$:2 nulls) using rigid termination with imbedded sensor. A series: termination was glued to end of tube. B series: termination was glued and surrounded with clay seal.

Trial	$Ch1$ (cm)			$Ch2$ (cm)	
~null frequency (kHz)	1.1	3.4	5.6	1.6	4.8
A-1	7.67	7.67	7.68	5.29	5.32
A-2	7.67	7.67	7.67	5.29	5.32
A-3	7.67	7.66	7.67	5.29	5.32
B-1	7.69	7.68	7.68	5.30	5.33
B-2	7.69	7.68	7.68	5.30	5.36
B-3	7.69	7.67	7.68	5.30	5.33

With the anechoic termination, the transfer function was measured for the sensors in their normal position, and also with the positions switched. The microphone spacing was then calculated using the phase delay, Φ, between the two calibrated sensors, normalized with respect to frequency, as defined in Eq. (6):

$$s = -\frac{\Phi(H_{12}/\sqrt{H_{12}\times H_{21}})}{2\pi f}. \tag{6}$$

As the termination is not completely anechoic at low frequencies, there is some variation/oscillation in the result as a function of frequency. The results of the separation calculation are shown in Fig. 5. To determine a single spacing distance, an average of the separation values was determined. The average value over 1–6 kHz was used. The wide variations at lower frequencies cannot be used with confidence.

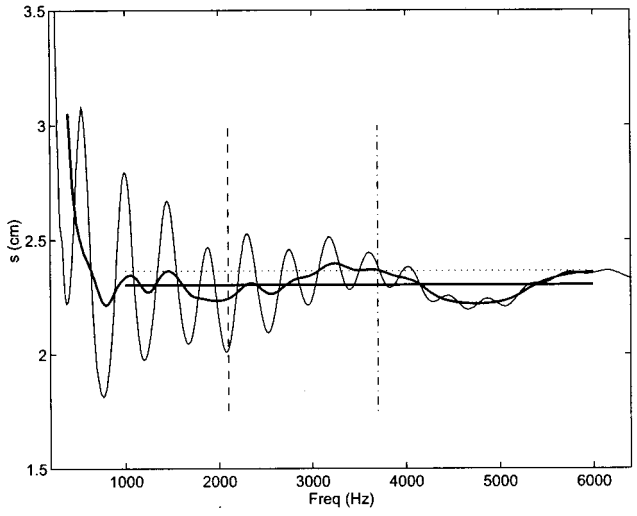


FIG. 5. Microphone separation distance calculation determined from measured phase delay. Thick curve shows a running average. The horizontal lines correspond to the (———) average value from 1 to 6 kHz and (···) distance determined using the frequency null method. Vertical lines correspond to (––) $ka=1$ and (–·–) the optimal performance frequency for the microphone spacing used.

The optimum frequency for the given microphone spacing which is equivalent to a quarter wavelength, $f_{\text{opt}}=c/4s=3.7$ kHz, and the location of $ka=1$ are also indicated. Note that at this frequency the two methods agree.

Results of this measurement are a separation distance of 2.30 cm, with a repeatability of ± 0.006 cm, within the measurement standard. In comparison to the frequency null method for measuring the distance devised above, though, there is a definite difference on the order of 0.06 cm, which is above the standard's threshold for determining the distances. The exact nature of the difference between the measurement techniques is unclear. For a typical measurement, accuracy of 0.06 cm would probably be sufficient. In this case, it is unclear how one would go about measuring an acoustic distance to any greater accuracy and be confident in the results. The results from the new method for distance measurements, using the frequency null interpolation, were chosen to be used in subsequent calculations as both s and l are required, and the anechoic spacing measurement only determines s. The resulting distances used in measurements therefore are $s=2.37$ cm and $l=5.31$ cm.

IV. RESULTS

A. Reference terminations

As there are no materials which have an absolutely defined reflection coefficient or impedance, there are limited tests which can be performed to verify the calculations and assumptions. The case of an infinitely rigid termination is theoretically possible, but in all practical purposes cannot be achieved or measured experimentally due to electrical noise, losses through the tube walls, or other factors. Measurements of a rigid termination, though, aid in showing the limitations of the specific experimental apparatus. The same holds true for an anechoic termination. It is impossible to make a totally anechoic termination, but an experimental effort towards such can be measured and used to further quantify the behavior of the measurement system.

The impedance of an ideally rigid termination would be infinite. It is therefore not practical to compare the measured impedance to the theoretical value. On the other hand, the reflection coefficient is well defined and is suitable to use for these evaluations. An ideal rigid termination has a reflection coefficient $|R|=1$. An ideal anechoic termination has a reflection coefficient $|R|=0$. The experimental measurements of the physical implementation of these idealized terminations are shown in Fig. 6. There are some losses in the system since the tube is not infinitely rigid, as seen in the low-frequency limit of the rigid termination. Additional errors can be seen in the low-frequency limit of the anechoic termination, which, when considered with the rigid and duct size limit, restrict the functional frequency range of the system to 1–6 kHz. Over this frequency range, the results of this measurement for the rigid termination produce a mean reflection coefficient of $\overline{|R|}=0.97$, which for all practical purposes can be considered rigid. The anechoic termination measurement results in a mean reflection coefficient of $\overline{|R|}=0.05$ over this range. For all practical purposes this anechoic termination is sufficient.

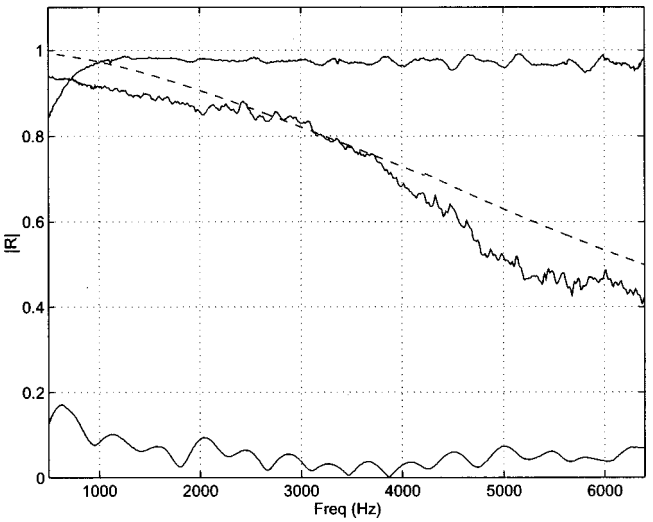


FIG. 6. Reflection coefficient reference measurements showing the measured values for a rigid (top), open (middle), and anechoic (bottom) terminations. Theoretical values for rigid=1 and anechoic=0. Theoretical value for the open termination is shown by the dashed line.

The only termination which has a well-defined solution, other than the idealized rigid and anechoic terminations, is an open-ended termination. The reflection coefficient for an unflanged thin-walled circular duct has the analytical approximation given in Eq. (7),

$$|R|=\begin{cases} e^{-(1/2)(ka)^2}\left(1+\frac{1}{6}(ka)^4\left[\ln\left(\frac{1}{\gamma ka}\right)+\frac{19}{12}\right]\right), & ka<1, \\ e^{-ka}\sqrt{\pi ka}\left(1+\frac{3}{32}\frac{1}{(ka)^2}\right), & ka>1, \end{cases} \tag{7}$$

where a is the tube cross-sectional radius and $\ln(\gamma)=0.5772$.[10] Comparison of the measured and theoretical value for the open ended tube termination can be seen in Fig. 6. The error in the theoretical approximation is less than 3% in the region between $ka=1$ and the optimal operating frequency of the design. The frequency at which there is the minimal amount of error in the measurement (as described in Sec. II B) is 3.7 kHz. This is the frequency at which the sensors are at λ/4. At this frequency, and down to approximately 2.2 kHz ($ka=1$), the measured response agrees very well with the theory.

The impedance for an anechoic termination is well defined. The impedance of a totally absorptive material should be simply the impedance of the propagating medium, in this case air. Plotting the specific acoustic impedance ratio, defined as the specific acoustic impedance normalized by the characteristic acoustic impedance of the medium ($\rho_0 c$), results in Fig. 7. The mean values for the anechoic impedance ratio measurement over the selected frequency range are $\overline{|Z|}=1.02$, $\overline{\text{Re}\{Z\}}=1.01$, and $\overline{\text{Im}\{Z\}}=-0.02$. Ideally, the real component of the impedance ratio should be equal to unity and the imaginary component equal to zero. These results are sufficiently close to the theoretical values.

Deviations from the theory below $ka=1$ could be due to low-frequency limitations of the two-microphone technique, which occur when the phase difference between the sensor

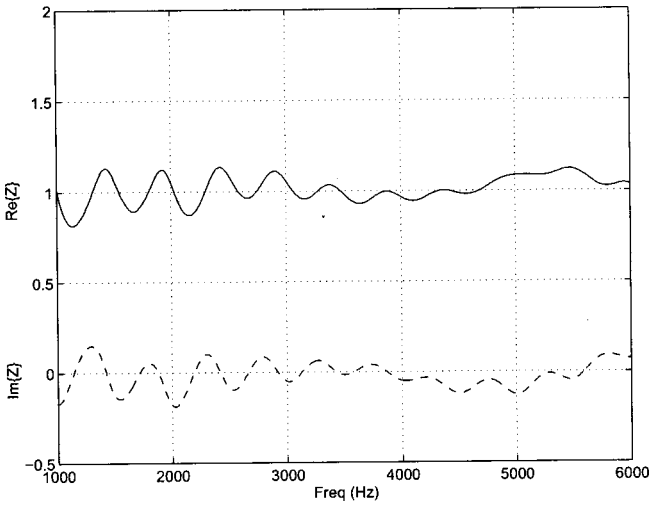


FIG. 7. Impedance measurement of the anechoic termination. The specific acoustic impedance ratio Z is plotted as $Z_s/\rho_0 c$ with (———) Re $\{Z\}$ and (–·–) Im $\{Z\}$. An ideally anechoic termination would have Re $\{Z\}=1$ and Im $\{Z\}=0$.

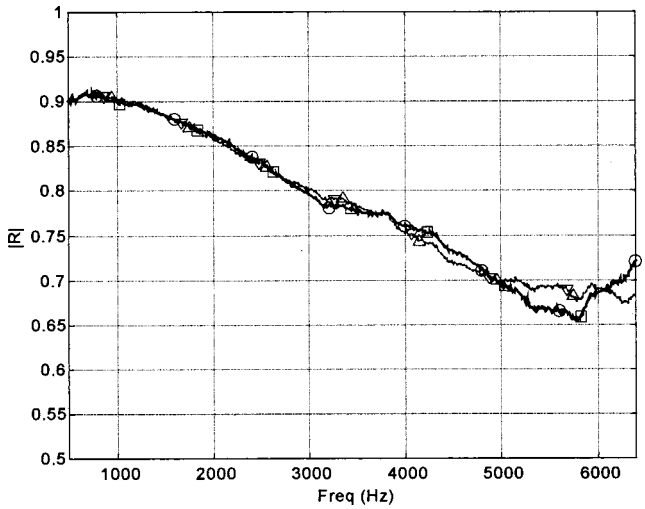


FIG. 8. Variations in the measured reflection coefficient of a 1.5-cm dense foam sample for various microphone location values: (○) $s=2.37$ and $l=5.31$, (▽) $s=2.30$ and $l=5.2$, (□) $s=2.30$ and $l=5.31$, and (△) $s=2.37$ and $l=5.2$.

locations is too small. Increased error above the designed optimum frequency is expected to some degree. The amount of deviation for this measurement could be due to the inaccuracy of the idealized theory to take into account the wall thickness of the tube at these higher frequencies, as well as the condition of $s \ll c/(2f_{\max})$ becoming less applicable. It is expected that over the range where the measurement corresponds to the theory, the error in further measurements will be small, increasing as the range is exceeded. From the results of the reference measurements of the rigid, anechoic, and open terminations there can be confidence in the measurement of the unknown acoustical properties of future test samples.

B. Test sample

A test sample was used to examine the effects of the new method for determining the microphone locations in the two-microphone method. A 1.5 cm thick disc of dense acoustic foam was measured. The variation on the acoustic absorption and specific acoustic impedance was evaluated for different microphone location estimations. Two possibilities for each distance were utilized in this investigation. The distance between the microphones, s, can be either 2.30 cm as determined using the phase method or 2.37 cm from the new frequency null method. The distance to the sample, l, can be either 5.2 cm as designed and measured by eye or 5.31 cm as determined using the frequency null method.

Using the same measured data transfer functions, the calculations for the reflection coefficient, R, and the specific acoustic impedance ratio, Z, were performed. As the sample is not a very rigid material, following Bodén and Åbom,[1] it is not expected that the effects due to distance errors would be noticeable. As can be seen in Fig. 8, there is not a great deal of variation in the calculation of the magnitude of the reflection coefficient due to errors in s and l. The discrepancy is about 3% between the new method and the old. An error of this sort can easily be tolerated in most experiments.

The variation in the calculation of acoustic impedance though is noticeably greater. As shown in Fig. 9 the errors between the methods are on the order of 12%. The results for the Re $\{Z\}$ vary quite substantially with $\Delta s=0.06$ cm and $\Delta l=0.11$ cm. This is a more serious difference and cannot be as easily ignored. As determined by Bodén and Åbom, these errors increase as the material sample becomes more rigid, where both the reflection coefficient and specific acoustic impedance increase.

V. CONCLUSION

In this article an improvement to the two-microphone impedance tube method for measuring acoustic absorption and specific acoustic impedance characteristics of a planar sample has been presented. This technique requires very accurate knowledge of the location of the sensors and sample

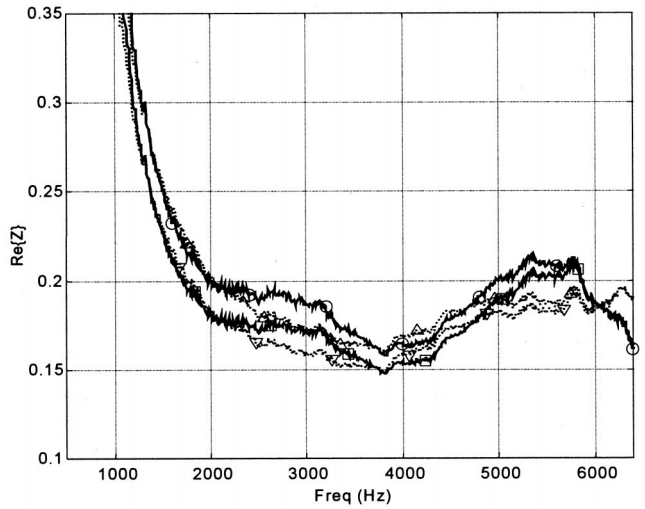


FIG. 9. Variations in the measured specific acoustic impedance ratio of a 1.5-cm dense foam sample for various microphone location values: (○) $s=2.37$ and $l=5.31$, (▽) $s=2.30$ and $l=5.2$, (□) $s=2.30$ and $l=5.31$, and (△) $s=2.37$ and $l=5.2$.

to a detail which exceeds visual and other nonacoustic measurement methods. This is due in large part to the fact that the required precision is smaller than the typical sensor microphone diaphragm, requiring an estimation of the location of the acoustic center of the microphone.

A new method is devised which utilizes a termination sensor and the measured transfer function between it and the two measurement microphones. From this measurement, the location of the sensors relative to the termination can be determined by calculating the frequencies at which the sensors are at a null. Results of this measurement method are compared to other methods, including visual observation and calculation of the microphone separation using phase delay measurements. The results of the new frequency null method do not agree with the previous methods within the measurement precision. But, the fact that this method produces both the microphone separation and the distance to the sample, while the phase delay method cannot produce the latter, is a marked improvement. In addition, the phase delay method shows a great deal of variation as a function of frequency which the new method does not.

The effects of an erroneous determination of the sensor locations was examined for a test sample of typical acoustic dense foam. Variation in the calculated reflection or absorption coefficient were on the order of 3% for induced errors of 2%–3% in the microphone positions, and is probably negligible with other experimental factors. Variations in the calculated specific acoustic impedance were greater, on the order of 12% for the same induced errors and test sample. This level of error is rather large for an unknown experimental error and must be considered if the two-microphone measurement method is being used for impedance measurements.

The errors due to these effects increase with the rigidity of the test material, making reliable measurements of high impedance materials a nontrivial task. Finally, the ASTM standard for this measurement technique requires a precision which exceeds the variations between the two methods, making the choice of distance calculations very important, the effects of which must be examined when making these sorts of acoustic measurements to ensure that accurate and reliable results are obtained. It is also questioned whether or not the standard is reasonable, as it requires precision in excess of most measurement techniques.

ACKNOWLEDGMENTS

The author is grateful to his supervisor Dr. Jiri Tichy for his assistance in this work. Further thanks are offered to Timothy Leishman for his assistance with the impedance tube and providing insight and encouragement throughout this work. Part of this work was financed by the Applied Research Lab and the Acoustics Department of The Pennsylvania State University.

[1] H. Bodén and M. Åbom, ''Influence of errors on the two-microphone method for measuring acoustic properties in ducts,'' J. Acoust. Soc. Am. **79**, 541–549 (1986).
[2] W. T. Chu, ''Transfer function technique for impedance and absorption measurements in an impedance tube using a single microphone,'' J. Acoust. Soc. Am. **80**, 555–560 (1986).
[3] J. Y. Chung and D. A. Blaser, ''Transfer function method of measuring in-duct acoustic properties. I. Theory,'' J. Acoust. Soc. Am. **68**, 907–913 (1980).
[4] J. Y. Chung and D. A. Blaser, ''Transfer function method of measuring in-duct acoustic properties. II. Experiment,'' J. Acoust. Soc. Am. **68**, 914–921 (1980).
[5] A. F. Seybert and D. F. Ross, ''Experimental determination of acoustic properties using a two-microphone random excitation technique,'' J. Acoust. Soc. Am. **61**, 1362–1370 (1977).
[6] C. Suzuki, H. Yano, and H. Tachibana, ''A new method for measuring normal incident sound absorption characteristics of materials using acoustic tube,'' J. Acoust. Soc. Jpn. (E) **2**(3), 161–167 (1981).
[7] ASTM E 1050-90, ''Standard Test Method for Impedance and Absorption of Acoustical Materials Using a Tube, Two Microphones, and a Digital Frequency Analysis System.''
[8] A. D. Pierce, *Acoustics* (Acoustical Society of America, New York, 1991).
[9] B. F. G. Katz, ''Measurement and Calculation of Individual Head-Related Transfer Functions Using a Boundary Element Model Including the Measurement and Effect of Skin and Hair Impedance,'' Doctoral thesis, The Pennsylvania State University, 1998.
[10] H. Levin and J. Schwinger, ''On the radiation of sound from an unflanged circular pipe,'' Phys. Rev. **73**(4), 383–406 (1948).

Acoustic absorption measurement of human hair and skin within the audible frequency range

Brian F. G. Katz[a)]
Acoustics Department, The Pennsylvania State University, State College, Pennsylvania 16801

(Received 5 January 1999; revised 18 September 1999; revised 17 May 2000; accepted 7 August 2000)

Utilizing the two-microphone impedance tube method, the acoustic absorption of human skin and hair is measured in the frequency range 1–6 kHz. Various locations on a number of human subjects are measured to determine if the presence of bone or an air pocket affects the acoustic absorption of human skin. The absorption coefficient of human hair is also measured. Additional techniques are utilized to minimize errors due to sample mounting methods. Techniques are employed to minimize potential errors in sensor and sample locations. The results of these measurements are compared to relevant historical papers on similar investigations. Results for skin measurements compare well with previous work. Measured hair absorption data do not agree with previous work in the area but do coincide with expected trends, which previous works do not. © *2000 Acoustical Society of America.* [S0001-4966(00)02411-5]

PACS numbers: 43.58.Bh, 43.55.Ev, 43.20.Ye, 43.80.Cs [SLE]

I. INTRODUCTION

There have been relatively few research efforts investigating and quantifying the acoustical properties of skin and hair over the audible frequency range. The majority of work regarding skin tissue measurements has been in the ultrasonic range. A number of mechanical studies have been done in which the mechanical impedances of various bone structures have been tested using impact hammers.[1]

The primary work in the area of acoustical measurements of skin and hair impedances was done in the early 1950's. Skin measurements were performed on the upper arm and thigh of human subjects using a plane-wave impedance tube and plane-wave excited rod.[2] The experimental design analyzed the change in the fundamental mode resonance frequency due to changes in surface impedance. A plane-wave tube was used for low frequencies (100–1500 Hz) while a rod was used, using the same technique, for higher frequencies (1.5–20 kHz). Results showed absorption coefficients, α, ranging from 0.04–0.004. These results were of the same order as the mechanical impedance results determined by Franke.[1]

Additional measurements of skin and hair utilized normal ''haired'' and hairless mice.[3,4] Two techniques were used. This first was to place the test animal very near to a high powered acoustic source (a 162-dB siren was used). The amount of sound absorption was calculated by measuring the increase in body temperature of the animal. This technique was used over a frequency range of 6–22 kHz. This measurement lasted for about an hour, by which time the animal died (another part of the experiment). From these measurements, the hairless mice had an absorption coefficient of 0.01–0.13 over the frequency range where the amount of absorption increased with frequency. In contrast, the haired mice had values ranging from $\alpha=0.5-1.1$ over the same frequencies. A second method utilized a reverberant chamber, over the frequency range of 1–22 kHz, in which the variations in the reverberation time were calculated due to the presence of the mice. Hairless mice had absorption coefficients ranging from $\alpha=0.14$ to $\alpha=0.16$ while haired mice had absorption coefficients starting at $\alpha=0.40$, increasing to $\alpha=0.95$ at 22 kHz.

This work is part of a larger project to gain a better understanding of the human hearing mechanism and the Head-Related Transfer Function (HRTF).[5] As part of that goal, we wish to examine whether or not the acoustical properties of skin and hair could contribute to the HRTF. To further this work it is necessary to have realistic and applicable measured data for the absorption of human skin and hair within the audible frequency range. Historical measurements of these parameters are limited, as described above. The previous work on the measurement of skin impedance was seen to be somewhat limited due to the usage of a mechanical excitation of the skin which may not have responded correctly for use in this airborne model. In addition, the technique had possible errors due to changes in the tension and curvature of the skin under test conditions. The previous work on hair, through the use of mice, does not seem adequate for use as a substitute for human hair. This is highlighted by the variations reported for different animal types, such as guinea pig and rat.[3,4] This work focuses on the measurement of the normal incident acoustic absorption of human skin and hair.

II. THEORY

The measurement method used to determine the absorption values utilized a two-microphone impedance tube technique.[5–12] This system provides for the measurement of the reflection coefficient and complex impedance of a material sample. Several assumptions and limitations exist regarding this method and shall be explored. These conditions

[a)]Currently with Arup Acoustics, 155 Avenue of the Americas, New York, NY 10013.

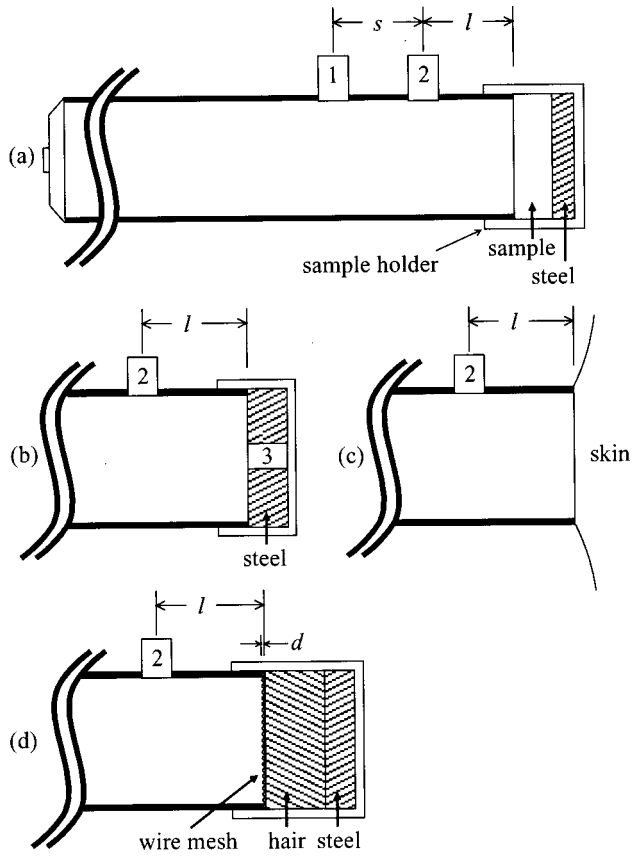


FIG. 1. Impedance tube measurement apparatus configuration. (a) Standard tube setup showing *Ch*1 and *Ch*2 measurement sensors. (b) Tube setup with termination sensor *Ch*3 for use in measurements involving sensor distances (Ref. 11). (c) Setup during skin measurement; note how the tube is pulled slightly away from the normal skin position. (d) Setup with hair sample and wire mesh in place; note that l is increased by the mesh insertion distance d so that l is the distance from *Ch*2 to the surface of the material.

include the assumption of a planar surface, normal incidence, and the accurate estimation of the distance from the microphones to the sample. The issue of accurate microphone location and sample distance was addressed by Katz and his methods are utilized here.[11]

III. TECHNIQUE

A. Apparatus

The parameters of the system, tube diameter, and sensor locations were designed to obtain optimum results over the frequency range of interest. The frequency range of interest was 1–6 kHz, the center frequency being 3.5 kHz. This frequency range was a function of the limitations of future models in which the data acquired were to be used. In addition, the measurement of higher frequencies using this method requires smaller diameter tubes and closer microphone spacings, introducing a variety of potential errors.

The apparatus for this experiment is identical to that used by Katz. A detailed description of the theory and setup can be found in the relevant references.[5,11] Several terminations were constructed to test and calibrate the system, in addition to the experimental sample terminations. A diagram of key components of the apparatus relevant to these measurements is shown in Fig. 1.

TABLE I. Microphone distance, l, for each frequency null measured (*Ch*1:3 nulls, *Ch*2:2 nulls). Results for *Ch*1 equal distance from *Ch*1 to the termination $(l+s)$ minus microphone spacing distance, (s), of=2.37 cm. Data labels: *gn*: termination glued to end of tube. $g+cn$: termination glued and clay around seal. $g+c+$screen n: termination glued with screen glued in place and clay around seal.

Distance data	*Ch*1 (cm) $((l+s)-s)$			*Ch*2 (cm) (l)	
~Null frequency (kHz)	1.1	3.4	5.6	1.6	4.8
$g1$	5.30	5.30	5.31	5.29	5.32
$g2$	5.30	5.30	5.30	5.29	5.32
$g3$	5.30	5.29	5.30	5.29	5.32
$g+c1$	5.32	5.31	5.31	5.30	5.33
$g+c2$	5.32	5.31	5.31	5.30	5.36
$g+c3$	5.32	5.30	5.31	5.30	5.33
$g+c1+$screen 1	5.32	5.44	5.45	5.36	5.45
$g+c2+$screen 2	5.31	5.44	5.45	5.36	5.45
$g+c3+$screen 3	5.32	5.44	5.45	5.36	5.45

According to the ASTM standard for the two-microphone impedance tube measurement, the distance from the end microphone to the surface sample must be known to an accuracy of 0.1 mm.[13] In the standard this requirement is waived for highly absorptive, roughly textured surfaces. This could apply to hair, but definitely not to skin measurements. For calibration with a rigid termination, and for skin which is assumed to be reasonably rigid at this point, this exception does not hold. Therefore, it is necessary to determine the distance from this microphone to the rigid or semirigid termination, as well as the microphone separation distance.

Employing the frequency null interpolation method devised by Katz,[5,11] several trials were performed and various termination conditions were used to determine these distances. The resulting distance corresponds to a measured frequency null in the transfer function between the measurement microphone and the reference microphone, *Ch*3, located at the termination end of the impedance tube [see Fig. 1(c)].

As the hair measurements required some means of being contained within a fixed area of the measurement system, a coarse wire mesh was used to keep the hair from moving, and was backed by the rigid termination and end cap [see Fig. 1(d)]. The mesh had a wire diameter of 0.24 mm with mesh hole size on the order of 1 mm^2. The distance between the microphones, s, was determined in the previous paper to be 2.37 cm.[11] The distance to the rigid sample was repeated with the screen in place to accurately determine the new value for l following the procedure designed by Katz.[11] Using the final set of data in Table I, the revised distance to the termination with the inclusion of the screen can be determined.

Various sealing methods for the sample container were used to determine the presence of any leakage. A small leak or gap could result in an error in the acoustical distance to the termination as well as the measured absorption. Effects of the various sealing methods can be seen in Table I. The results shown in Table I present consistent results over frequency and sealing methods.

To ensure that the screen caused minimal acoustic interference with the measurements, the reference termination measurements were repeated with the screen in place. As

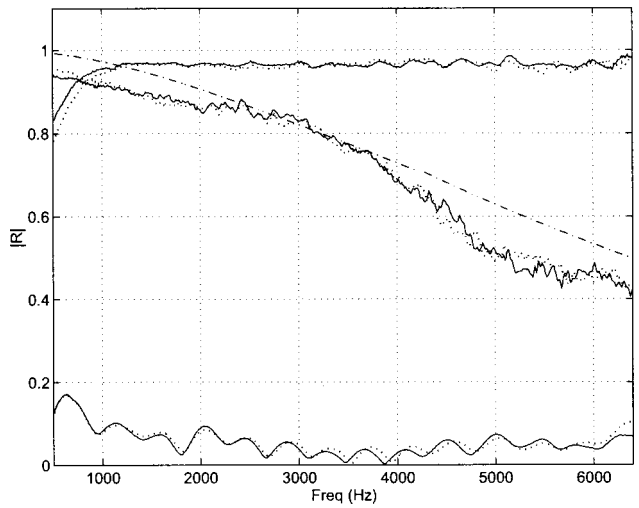


FIG. 2. Reflection coefficient reference measurements showing the measured values for rigid (top), open (middle), and anechoic (bottom) terminations with (···) and without (——) wire screen mesh in place. Theoretical values for rigid ($|R|=1$) and anechoic ($|R|=0$). Theoretical value for the open termination is shown by the dashed line (Ref. 16).

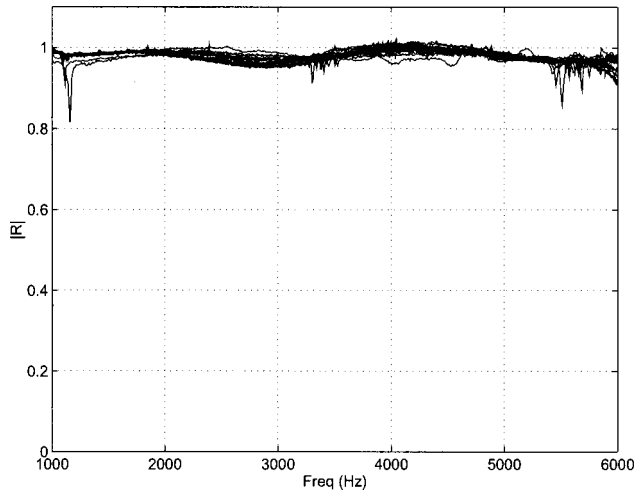


FIG. 3. Reflection coefficient measurement for a number of different skin locations and conditions.

shown in Fig. 2, inclusion of the screen has no effect on the measured reflection coefficient of the rigid termination, the anechoic termination, or the open termination.

It is interesting to note that the inserted distance increases with frequency. The high-frequency nulls measured show a screen distance insertion of 0.15 cm, the mid-frequency nulls show a distance of 0.12 cm, and the first and lowest frequency nulls show almost no effect on distance: 0.0–0.05 cm. It is difficult to determine which result to use, and how to interpret this discrepancy. It may throw possible doubt on the measurement scheme, or it may point out a low-frequency limitation of the method. The latter is more plausible reasoning as it has been shown in a number of places that the low-frequency response is not ideal (see Katz[11] and Fig. 2). It is also plausible that the acoustic effect of the wire mesh is frequency dependent. Measurement of the distance to account for this would be very difficult as there are variations in the measured distance due to other factors, such as losses in the tube and a nonideally anechoic termination, which would need to be removed before such a detailed measurement could be made. The higher frequency results will be used, giving a distance to termination value of $l=5.45$ cm with the presence of the screen.

B. Planar surface assumption

For semirigid to rigid impedance terminations the distance to the termination is crucial, as described in the previous section. Therefore, it is important that the sample be flat to roughly the same degree of accuracy as the standard requires. In measuring the properties of skin, the measurement tube was placed against the area of skin to be measured. It was quickly observed that a curvature of the skin was induced which increased with the pressure applied with the tube. As discussed in the literature, and as would be expected, the impedance characteristics of the skin will vary with the applied pressure as the tension on the area will vary, creating a different mechanical system. It would be difficult to create a measurement system that could reliably use the same applied force for each trial, simulate normal skin condition, and provide a good acoustic seal. The following procedure was therefore conceived to minimize these problems. The impedance tube was attached to the skin area with a thin layer of nontoxic rubber cement, providing a good acoustic seal. Once cured, the tube was pulled slightly so that the skin sample was visibly raised from the skin surface area as represented in Fig. 1(c). The result of this procedure was that the skin sample was under little tension in the plane of the skin, keeping close to the normal skin position situation, without effects due to the applied pressure of the tube. Having the skin slightly pulled away from the normal surface removed the effect of bowing or curvature due to the pressure of the tube against the skin, and any normal curvature present on the sample area. The result was a planar skin sample within the impedance tube with reasonably typical skin tension in the planar direction. It was acknowledged that the tension in the direction normal to the surface would not be the same as the typical position, but this was assumed to have much less of an effect than the other conditions the procedure was designed to alleviate.

IV. MEASUREMENTS

A. Skin

The main purpose of the acoustical measurement of the skin was to determine if skin behaves as a rigid material, or exhibits more complex behavior in the selected audible frequency range. To this aim, acoustic absorption measurements were made at the following locations on several live human subjects: forearm, upper arm, center of forehead, stomach (under last rib), cheek with teeth clamped (mouth closed), cheek with teeth loose (mouth closed), cheek with mouth open, and cheek puffed. The results for all these measurements are given in Fig. 3, showing relatively equal results for all test situations.

The average value for the reflection coefficient for these measurements was $\overline{|R|} \cong 0.97$, or $\overline{\alpha} \cong 0.03$. In comparisons

TABLE II. Hair physical properties including total mass, hair diameter, and density within test chamber of the indicated samples. Also included are data for foam samples (*kf* and *wf*) with approximate fiber diameters. All samples were placed in a test chamber with a volume of 7.69 cm^3. Variations of hair diameter within a given sample are also included.

Sample ID	Mass (g)	Fiber diameter (mm)	Bulk density (kg/m^3)
*h*1	1.5811	0.047±0.002	206
*h*2	1.0415	0.045±0.001	130
*h*3*a*	1.7204	0.055±0.01	233
*h*3*b*	0.6645	0.055±0.01	86
*h*4	1.9630	0.057±0.002	255
kf	0.2768	≈0.0508	36
wf	0.1590	≈0.0635	21

to previous works showing absorption coefficients ranging from 0.004 to 0.04, these results agree well.[1,2] A reflection coefficient of 1.0 is a characteristic of an ''acoustically rigid'' surface. For all practical purposes skin can be assumed to be rigid, as the reference measurement for the rigid termination also had the same result.

B. Hair

Measurements of the acoustical properties of hair were made on samples from four individuals. A fifth sample contained more hair from one of the first four individuals. In addition, two samples of two different acoustical foams were included for comparison. Details of the samples are given in Table II. Hair samples were taken from these subjects and placed in the impedance tube as shown in Fig. 1(d). The dimensions of the cylindrical sample chamber were as follows: length=1.5, diameter=2.60 cm. The bulk density of the various samples can be determined using the mass as given in Table II and a volume of 7.69 cm^3. The bulk density for hair sample *h*1 can be estimated at 206 kg/m^3, comparable to heavy rockwool material samples. The porosity of the hair samples was not measured at this stage of the experiment. Methods for measuring the flow resistance and compressibility of fibrous materials, such as hair, have been developed by Tarnow.[14] Approximation methods for determining these parameters for low frequency and low packing density have also been developed.[15]

Of initial interest was the variation in hair diameter from sample to sample, and especially the variation within a sample. Hairs *h*1, *h*2, and *h*4 have little variation in diameter within the sample, while *h*3 has a large variation (a factor of ×10 more variation). One explanation for this could be the method of hair acquisition. Samples *h*1 and *h*2 were taken from a single location on the head, grouping the hair in a small bunch and cutting the bunch with scissors. This could ensure a more uniform distribution. Samples *h*3 and *h*4 were taken from clippings from a haircut. It is possible that over the entire head (and neck) hair diameters vary and that these various hair types were included in *h*3. Perhaps there was less variability in *h*4 because of the individual or the type of hair which was cut.

Measured absorption coefficients for the hair samples are given in Fig. 4. Data for two foam material samples (*kf* and *wf*) are included for comparison. Figure 5 shows the absorption reported in the literature for various animals and

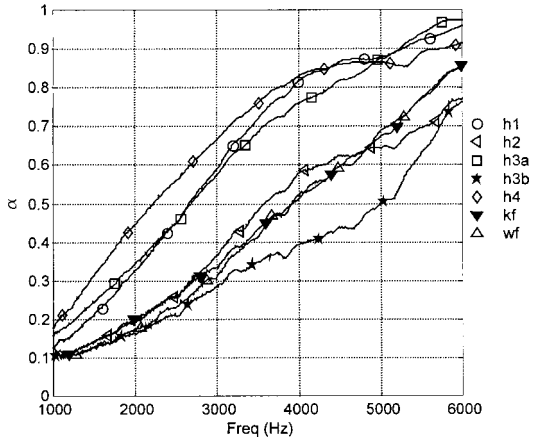


FIG. 4. Absorption coefficient measurements for hair and samples of two foams used in the construction of the anechoic termination.

fiberglass for comparison.[3,4] The amount and parameters of the hair on a mouse were not reported in the literature. It is quickly apparent from the results that hair is not a rigid surface. There is some diversity in the hair results, while the results for the foam are almost identical. Examining the physical properties of the samples (see Table II) shows that the reflection coefficient generally increases as the mass of the hair sample decreases. However, with similar fiber diameter and significantly less mass, the foam samples have comparable absorption. This discrepancy is attributed to the complex nature of fibrous absorption and its relation to fiber length, packing geometry, fiber material, etc. The results of this measurement vary from previous data, as shown in Fig. 5, for furred mammals over this frequency range, where α varied from 0.2–0.6. The absorption values were dependent upon the measurement method and animal tested (mouse, rat, or guinea pig).

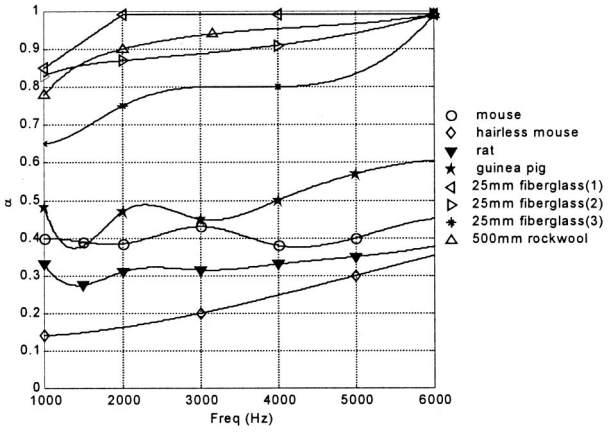


FIG. 5. Absorption coefficient measurements for normal mice, hairless mice, rat, and guinea pig samples using reverberation chamber technique. (Data for animals from Refs. 3 and 4.) Absorption coefficient measurements for 25-mm fiberglass on rigid backing. (1) Has a packing density of 60 kg/m^3, from Ref. 17. Data (2) from Ref. 18. (3) Has a packing density of 70–100 kg/m^3, from Ref. 19. Absorption coefficient measurement for rockwool with packing density of 150 kg/m^3 from Ref. 19.

V. CONCLUSION

In this paper measurements have been presented of the acoustic absorption of human skin and hair over the frequency range of 1–6 kHz. Measurements were made using the two-microphone impedance tube method. Refinements were used to accurately determine the distance between the microphones and the sample materials as presented by Katz.[11] In addition, test procedures were employed to minimize the error due to these distances and the nature of the samples.

The reliability of the measurements is justified by using several terminations with known absorption characteristics. These terminations consisted of a rigid steel disc, an anechoic wedge, and an open-ended tube. The response of the apparatus to these reference measurements shows very good agreement in the middle of the frequency range of interest, with results becoming less confident at the frequency extremes.

A variety of measurements were made on human skin, taking into account the possible effects of bone and air pockets beneath the skin. All measurements consistently showed that human skin is acoustically rigid, with $\alpha = 0.03$. The use of an adhesive and reverse tension on the skin provided for a very flat sample surface. This agreed well with previous works using alternate measurement schemes.

Absorption measurements of human hair samples were conducted for a number of subjects with various hair properties. The results were varied, though all followed similar trends. These trends correspond to the absorption behavior of typical foam or mineral wool. The variations in the results are most likely due to the variations in hair sample length and sample packing densities and are not out of the range of variation seen in typical published data. The results do not agree with previous work on mice and other furred animals which were designed to estimate the absorption of human hair.[4] Results of this study provide data, based on actual human test samples, which more nearly accurately characterize the acoustic absorption of human hair and skin in the audible frequency range.

ACKNOWLEDGMENTS

The author is grateful to his supervisor, Dr. Jiri Tichy, for his persistence in this work and to Dr. Courtney Burroughs for his participation in many discussions. Thanks are also offered to Mr. Timothy Leishman for his assistance with the impedance tube and providing insight and encouragement throughout this work. Part of this work was financed by the Applied Research Lab and the Acoustics Department of The Pennsylvania State University.

[1] E. K. Franke, ''Mechanical impedance measurement of the human body surface,'' Air Force Technical Report No. 6469, Aero Medical Laboratory, Air Material Command, Wright–Patterson Air Force Base, Dayton, Ohio, 1951.
[2] H. E. von Gierke, ''Measurement of the acoustic impedance and the acoustic absorption coefficient of the surface of the human body,'' Air Force Technical Report No. 6010, Aero Medical Laboratory, Air Material Command, Wright–Patterson Air Force Base, Dayton, Ohio, 1950.
[3] R. W. Farwell, ''Acoustic Absorption of Haired and Hairless Mice,'' Masters thesis, The Pennsylvania State University, 1955.
[4] ''Sound absorption at the surfaces of small laboratory animals,'' WADC Technical Report 57-461, Wright Air Development Center, June 1957.
[5] B. F. G. Katz, ''Measurement and calculation of individual Head-Related Transfer Functions using a boundary element model including the measurement and effect of skin and hair impedance,'' Ph.D. thesis, The Pennsylvania State University, 1998.
[6] H. Bodén and M. Åbom, ''Influence of errors on the two-microphone method for measuring acoustic properties in ducts,'' J. Acoust. Soc. Am. **79**, 541–549 (1986).
[7] W. T. Chu, ''Transfer function technique for impedance and absorption measurements in an impedance tube using a single microphone,'' J. Acoust. Soc. Am. **80**, 555–60 (1986).
[8] J. Y. Chung and D. A. Blaser, ''Transfer function method of measuring in-duct acoustic properties. I. Theory,'' J. Acoust. Soc. Am. **68**, 907–913 (1980).
[9] J. Y. Chung and D. A. Blaser, ''Transfer function method of measuring in-duct acoustic properties. II. Experiment,'' J. Acoust. Soc. Am. **68**, 914–921 (1980).
[10] B. F. G. Katz, ''New approach for obtaining individualized Head-Related Transfer Functions,'' J. Acoust. Soc. Am. **100**, 2609(A) (1996).
[11] B. F. G. Katz, ''Method to resolve microphone and sample location errors in the two-microphone duct measurement method,'' J. Acoust. Soc. Am. **108**, 2231–2237 (2000).
[12] A. F. Seybert and D. F. Ross, ''Experimental determination of acoustic properties using a two-microphone random excitation technique,'' J. Acoust. Soc. Am. **61**, 1362–70 (1977).
[13] ASTM E 1050–90, Standard test method for impedance and absorption of acoustical materials using a tube, two microphones, and a digital frequency analysis system.
[14] V. Tarnow, ''Measurement of sound propagation in glass wool,'' J. Acoust. Soc. Am. **97**, 2272–81 (1995).
[15] V. Tarnow, ''Airflow resistivity of models of fibrous acoustic materials,'' J. Acoust. Soc. Am. **100**, 3706–13 (1996).
[16] H. Levin and J. Schwinger, ''On the radiation of sound from an unflanged circular pipe,'' Phys. Rev. **73**, 383–406 (1948).
[17] D. B. Bies and C. H. Hansen, ''Sound absorption in enclosures,'' in *Encyclopedia of Acoustics*, edited by Malcolm J. Crocker (Wiley, New York, 1997).
[18] Lennart Karlen, *Akustik i rum och byggander* (Svensk Byggtjänst, 1983) (in Swedish).
[19] *Noise and Vibration Control*, edited by Leo Beranek and István Vér (Wiley, New York, 1992).

Measurements of anisotropic sound propagation in glass wool

Viggo Tarnow
Department of Applied Engineering Design and Production, Technical University of Denmark, Bygning 358, DK 2800 Lyngby, Denmark

(Received 30 November 1999; revised 9 August 2000; accepted 24 August 2000)

The attenuation coefficient and phase velocity of plane sound waves propagating in three perpendicular directions in glass wool were measured in the frequency range 50–10 000 Hz. For glass wool of mass density 14 kg/m^3 at the frequency 1000 Hz, the attenuation constant for propagation perpendicular to the glass wool sheets was 75 dB/m, and for propagation parallel with the sheets 57 dB/m. For mass density 30 kg/m^3, the corresponding numbers were 140 and 100 dB/m. The measured values were compared with calculated ones taking into account the movements of the fiber skeleton. The calculations need the elastic moduli, which were measured. For density 14 kg/m^3 and deformation perpendicular to the glass wool sheets, the static modulus was 2.0 kPa, and for parallel deformation 120 kPa. The corresponding numbers for mass density 30 kg/m^3 were 16 and 390 kPa. © *2000 Acoustical Society of America.* [S0001-4966(00)05611-3]

PACS numbers: 43.58.Vb, 43.20.Jr, 43.55.Ev, 43.35.Mr [SLE]

I. INTRODUCTION

Sound propagation in fiber materials such as glass wool has been studied intensely, but most studies have been done with plane waves propagating in only one direction with respect to the fiber orientation. Kirkby and Cummings[1] reported measurements of propagation constants and characteristic impedance for basalt wool and suggested methods of extrapolation to low frequencies, where the popular formulas of Delany and Bazley[2] are less reliable. Tarnow[3] measured the propagation constants and characteristic impedance of glass wools for one direction of sound propagation. Lambert[4] presented dynamic resistivity data for Kevlar fiber material, where a resonance at 850 Hz was found. Voronina[5] gave data for glass, mineral, and basalt wool. These authors used waves propagating in one direction only.

Bokor[6] reported measurements of sound attenuation and phase velocity in directions parallel with and perpendicular to fiberglass sheets for frequencies 400–1000 Hz. At 1000 Hz the parallel attenuation coefficient divided by the perpendicular one was 0.59, and the corresponding ratio between the real part of the wave numbers was 0.91. Allard, Bourdier, and L'Esperance[7] measured with oblique directions of sound propagation the impedance of glass wool mats placed on asphalt. They found the ratio 0.6 between direct current (dc) air flow resistivity parallel and perpendicular to the large surface of the glass wool mat, which had a perpendicular d.c. resistivity of 25 000 kg·s^{-1}·m^{-3}.

In this paper the measured sound attenuation coefficient and phase velocity for plane monochromatic sound waves propagating in three perpendicular directions will be presented for glass wool of densities 14 and 30 kg/m^3 in the frequency range 50–10 000 Hz. The measured attenuation coefficient and phase velocity will be compared with calculations, which use formulas of Wilson for the effective mass density and the compressibility.

II. CALCULATION OF ATTENUATION COEFFICIENT AND PHASE VELOCITY

Delany and Bazley[2] have developed formulas that can be used to compute the attenuation coefficient and phase velocity as a function of frequency. The derivation of these formulas was purely empirical, and they will not be used in the present paper, where we seek to identify the physical mechanism responsible for sound attenuation. The Delany and Bazley formulas do not provide for movements of the fiber skeleton and are not accurate at low frequencies.[1]

Fiber movement is significant, and was accounted for using a method developed previously.[8] The fiber skeleton can carry a stress wave that is coupled to the acoustic pressure wave. The air moves, and the viscous drag on the fibers makes them move. The complex pressure p of a monochromatic plane wave propagating through a layer of glass wool of thickness x can than be computed from

$$p = A(e^{ik_1 x} + Fe^{ik_2 x}), \tag{1}$$

where A is an undetermined constant, which value is not used, as only rations of pressures for different values of x will be considered. There are two complex wave numbers k_1 and k_2, the first of which is mainly associated with the pressure wave in the air, and the second is mainly associated with the stress wave in the fiber skeleton. F is a complex coupling factor given below.

The wave numbers can be computed from the following determinant equation:

$$\begin{vmatrix} -\omega^2(\mu - d\rho) + k^2 c_{nn} - i\omega R', & k^2 dK + i\omega R' \\ k^2 dK + i\omega R', & -\omega^2 \rho + k^2 K - i\omega R' \end{vmatrix} = 0, \tag{2}$$

where ω is the angular frequency, μ is the mass density of glass wool, d is the volume density of fibers ($d = 1 - \phi$, ϕ is the porosity), ρ is the static mass density of air, k is the unknown wave vector sought, c_{nn} is the elastic modulus of the fiber skeleton for the direction of wave propagation considered. R' is a factor that takes care of the viscous drag of

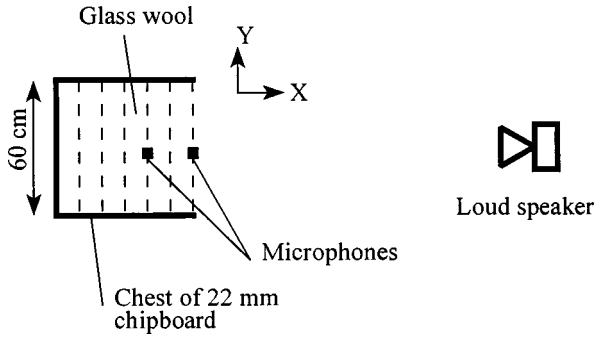


FIG. 1. The measurement setup in the anechoic room. Six slabs of glass wool each with dimension 10×60×90 cm were placed in the chest. The distance from chest to loudspeaker was 170 cm.

the air on the fibers, $R'=R+i\omega\rho$, where R is the complex resistivity to air flow through fixed fibers, $R=-i\omega\rho_e$, and ρ_e is the complex air mass density function, which is often used. The complex resistivity is preferred, because it is nearly constant and real for frequencies from d.c. to 1000 Hz. K is the bulk modulus of air, $K=1/\phi C$, where C is the complex compressibility of air itself.

The wave in the glass wool is excited by a plane wave in air that hits the plane surface of glass wool head-on at $x=0$. F was calculated from the boundary condition, zero stress, on the interface between air and glass wool, $x=0$. The details are described in the author's paper.[8] One has

$$F=\frac{(\omega R')^2}{(-\mu\omega^2+c_{nn}k_1^2-i\omega R')\cdot(-\rho\omega^2+k_2^2K-i\omega R')}. \quad (3)$$

To compute the sound pressure one has to know the complex air compressibility and resistivity of the glass wool with fixed fibers. Formulas for these were given by Attenborough.[9] They were derived by assuming that the glass wool consists of a solid with circular holes. The microstructure of real glass wool is far from this, and, therefore, they are not used.

Allard[10] proposed formulas for the compressibility and resistivity based on work by Johnson *et al.*,[11] who suggested a formula for interpolating the complex resistivity between static flow and very high frequencies. This formula depends on a parameter, the viscous length, defined from the very high frequency resistivity. It can be shown by direct computation that the suggestion of Johnson is not valid for a material consisting of parallel fibers. The viscous length can be computed for this model, and the complex resistivity can then be computed from Johnson's formula. But at frequencies about 2000 Hz, the real value of the resistivity computed from Johnson's formula becomes far smaller than the values computed by finite elements. This means that the Johnson procedure can not be used to compute attenuation of the glass wool samples with the densities considered in this paper; therefore, Allard's procedure is not useful for these glass wool samples.

We use the formulas proposed by Wilson[12] because they are simple, and do not make any assumption on the microstructure of the glass wool.

The air compressibility was calculated by Eq. (24) of Wilson,[12]

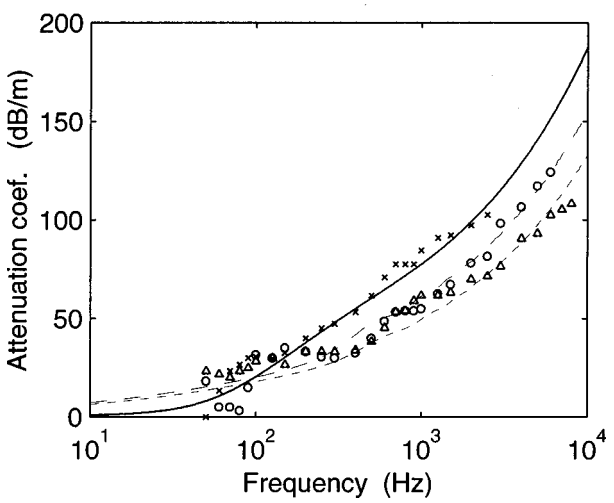


FIG. 2. Attenuation coefficient for glass wool of density 14 kg/m^3. The points marked ''x'' are the measured value for propagation in the *X*-direction. The points marked ''○'' are for the *Y*-direction, and ''△'' is for the *Z*-direction. The curves were calculated. The full line is for the *X*-direction, the dashed for *Y*-direction, and the dotted for *Z*-direction.

$$C=\frac{1}{\gamma P}\left[1+\frac{\gamma-1}{(1-i\omega\tau_c)^{1/2}}\right], \quad (4)$$

where $\gamma=1.40$ for atmospheric air, P static air pressure, and τ_c is a time constant, that was found from the measured compressibility in Ref. 13.

The resistivity can be calculated by Eq. (23) of Wilson[12]

$$R=\frac{-i\omega\rho}{1-(1-i\omega\tau)^{-1/2}}, \quad (5)$$

where ρ is the static mass density of air (1.20 kg/m^3 at 1 atm and temperature 20 °C). τ is a time constant, which in this paper was found by fitting of the calculated attenuations at high frequencies to the experimental ones. The resistivity calculated from Eq. (5) approaches the value $2\rho/\tau$ asymptotically at low frequencies.

III. MEASUREMENT METHOD AND RESULTS

Measurements were done in an anechoic room. Six glass wool slabs each measuring 10×60×90 cm were placed in a rectangular chest, made of 22-mm-thick chipboard and measuring 60×60×90 cm inside. There were no air spaces between the different sheets of glass wool. One side of the chest was open facing a loudspeaker 1.7 m away as shown in Fig. 1. In this way the glass wool pile was exposed to approximately plane sound waves from the loudspeaker. Three different sides of the chest could be opened, so sound waves in three different directions could be sent into the pile of glass wool. The sound pressure was measured with one microphone placed at the open surface of the glass wool and a second inside the glass wool. The sound pressures and their phases were recorded. The attenuation coefficient was found from the attenuation of the sound pressure in dB divided by the thickness of the glass wool between microphones. The phase velocity c was found from the measured phase shift ϕ and the distance between the microphones x. Thus $c=\omega x/\phi$.

TABLE I. Measured elastic moduli.

Direction	Density 14 kg/m^3	Density 30 kg/m^3
X	2.0 kPa	16 kPa
Y	120 kPa	390 kPa
Z	150 kPa	270 kPa

TABLE II. Time constants for complex compressibility and resistivity.

	Density 14 kg/m^3	Density 30 kg/m^3
Compressibility	0.50 ms	0.30 ms
Resistivity		
X-direction	0.36 ms	0.16 ms
Y-direction	0.55 ms	0.21 ms
Z-direction	0.80 ms	0.23 ms

The glass wool is manufactured as slabs with dimension $10\times60\times90$ cm. The direction of the 10-cm edge is called the *X*-direction, the *Y*-direction is the direction of the 60-cm edge, and the *Z*-direction is in the direction of the 90-cm edge. The glass fibers are mostly oriented in the *YZ*-plane, and their directions approximately randomly distributed, but the elastic moduli data seem to show that they are not completely randomly distributed. Measurements were done on two types of glass wool, having density 14 and 30 kg/m^3. Both types have a fiber diameter of 6.8 μm.

Figure 2 shows the result of measurements of attenuation coefficient of sound waves in glass wool of density 14 kg/m^3. Sound pressure was measured at the free surface and in the glass wool 30 cm from the free surface. The points on Fig. 2 for propagation in the *X*-direction are marked ''x.'' Over most of the frequency range, the attenuation coefficient is highest in this direction. The measurement uncertainty of the attenuation coefficient is 1.6 dB/m.

To compute the curves one needs experimental data. Tortuosity and porosity are assumed to be 1, because the most dense glass wool, 30 kg/m^3, has a relative fiber volume density of 0.016.

The elastic moduli of the glass wool samples were measured by a static method described in the author's paper.[8] The measured elastic moduli are presented in Table I. The glass wool is much softer in the *X*-direction than in the other directions.

The compressibility was computed from Eq. (4) with the time constant 0.50 ms, which was found from earlier measurements of the compressibility.[13] These measurements were done on glass wool samples with the same fiber diameters but a little higher density, 16 and 40 kg/m^3, where the samples studied here have 14 and 30 kg/m^3. Measurements of compressibility are very difficult, and the data are very scattered. Figure 3 shows the experimental data normalized by division with the isentropic compressibility. The full line curve was computed with time constants 0.50 ms. From the scattering of the data one can estimate the accuracy of the time constants to 20%. Fortunately the computation of attenuation of sound wave is not very sensitive to the value of these time constants. The time constant for the two glass wool samples is given in Table II.

The resistivity was computed from Eq. (5). The values of the time constants were found by fitting graphically the curves to attenuation data at high frequencies. The time constants are given in Table II. The attenuation is sensitive to the value of the time constant in contrast to the phase velocity, which is less sensitive.

The calculated attenuation coefficient shown in Fig. 2 was computed from Eq. (1) in the same way as it was measured. The value of the constant *A* in Eq. (1) has no influence on the attenuation per meter, and it was set to 1. The complex sound pressure was computed at the surface of the glass

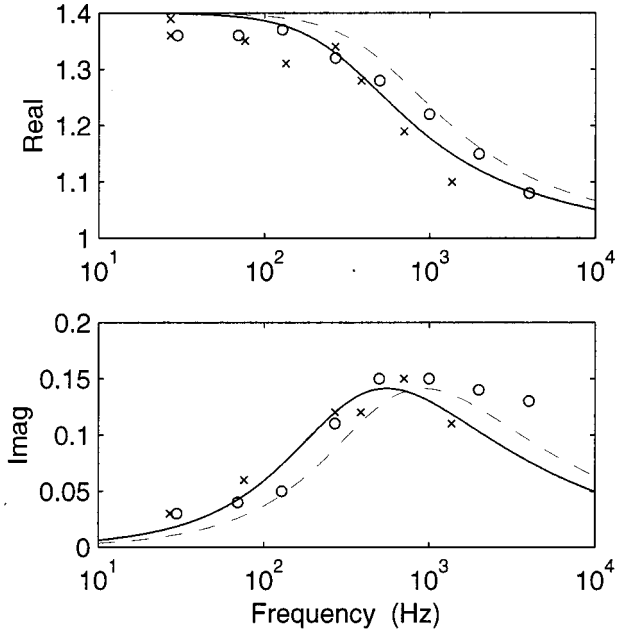


FIG. 3. The complex compressibility of air between fibers. The data are normalized by division with the isentropic compressibility. The ''x'' points are for density 16 kg/m^3, and the full line is computed by Eq. (4) with time constant 0.50 ms. The ''○'' points and the dashed line are for density 40 kg/m^3 with time constant 0.30 ms.

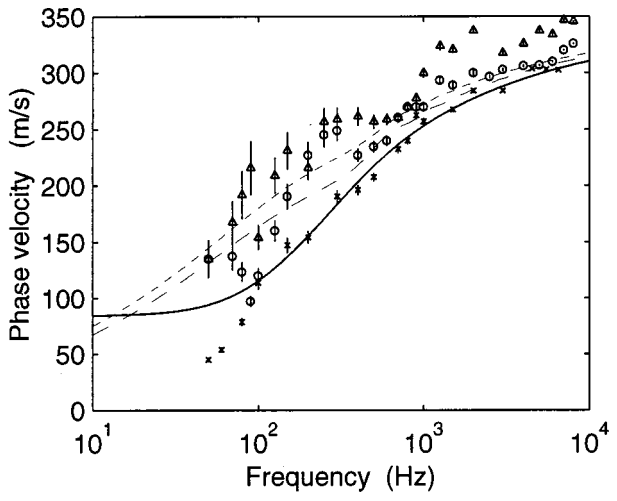


FIG. 4. Phase velocity for glass wool of density 14 kg/m^3. The points marked ''x'' are the measured value for propagation in the *X*-direction. The points marked ''○'' are for the *Y*-direction, and ''△'' is for the *Z*-direction. Measurement uncertainty is shown as vertical error bars. The curves were calculated. The full line is for the *X*-direction, the dashed for *Y*-direction, and the dotted for *Z*-direction.

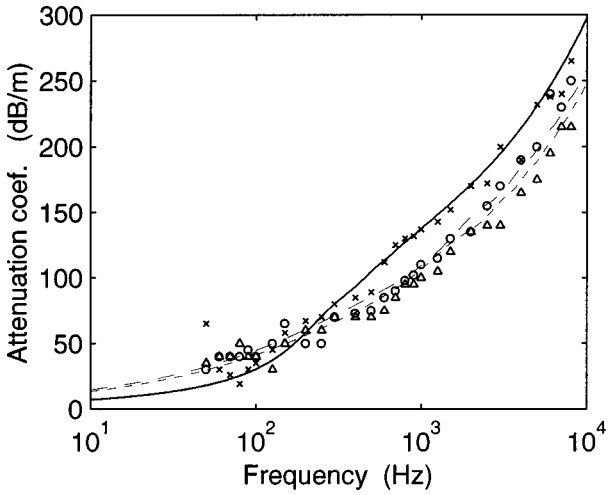


FIG. 5. Attenuation coefficient for glass wool of density 30 kg/m^3. The points marked ''x'' are the measured value for propagation in the *X*-direction. The points marked ''○'' are for the *Y*-direction, and ''△'' is for the *Z*-direction. The curves were calculated. The full line is for the *X*-direction, the dashed for *Y*-direction, and the dotted for *Z*-direction.

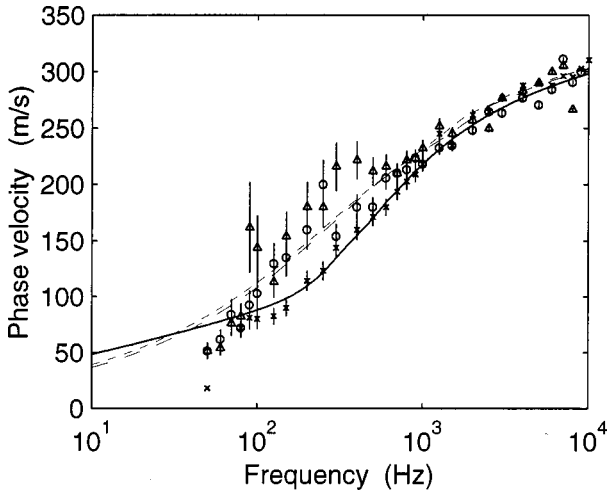


FIG. 6. Phase velocity for glass wool of density 30 kg/m^3. The points marked ''x'' are the measured value for propagation in the *X*-direction. The points marked ''○'' are for the *Y*-direction, and ''△'' is for the *Z*-direction. Measurement uncertainty is shown as vertical error bars. The curves were calculated. The full line is for the *X*-direction, the dashed for *Y*-direction, and the dotted for *Z*-direction.

wool ($x=0$) and 30 cm below the surface ($x=30$ cm). From the two amplitudes, the sound attenuation in dB was found and divided by 30 cm to give the attenuation coefficient.

Due to the two wave numbers in Eq. (1) the attenuation coefficient, which is the attenuation per length, is not strictly a constant but depends on the distance in the glass wool. However, the attenuation coefficient was also measured for thicknesses 10, 20 cm, and no dependence on thickness was detected. This was supported by calculations which showed deviations small compared to the experimental scattering of data. This means that for thicknesses 0–30 cm, one can regard the attenuation coefficient as a constant in the frequency range 50–10 000 Hz. It is tempting to neglect the movement of the fiber skeleton, but this will give a high attenuation coefficient in the *X*-direction at frequencies about 100 Hz.

At frequencies lower than 100 Hz, the measured points deviate from the calculated curves, because the sound pressure wave, which is reflected from the rear of the glass wool pile, interferes with the direct wave. This is not so at higher frequencies, because the reflected wave is much attenuated before it reaches the microphone inside the glass wool.

Figure 4 shows the phase velocity for glass wool of density 14 kg/m^3, measured with a glass wool layer of thickness 30 cm. The points marked ''x'' are for propagation in the *X*-direction. The curves were calculated from the same time constants that were used for calculating the attenuation. At frequencies lower than 100 Hz, the measured phase velocity is smaller than the calculated curve. This is due to reflection of waves from the rear. For propagation in the *Y*- and *Z*-direction at low frequencies, the measurement uncertainty of the phase angles gives a great uncertainty of the phase velocity, because the phase shift is small. The uncertainty is shown as vertical error bars on the graph. The high phase velocity for the *Z*-direction (triangular point) in the frequency interval 1000–10 000 Hz is difficult to explain, perhaps the glass wool is not homogenous on the scale of the wavelength. The phase velocity was also measured and computed for thicknesses 10, 20 cm. No dependence of phase velocity on thickness was found.

Figure 5 shows the measured attenuation coefficients for glass wool of density 30 kg/m^3, measured with 10-cm thickness of glass wool. This thickness was chosen because a thicker layer would attenuate the high frequency wave too much. The measurement uncertainty is 5 dB/m. The attenuation was also measured for thicknesses 20, 30 cm, and it was found to be constant.

Figure 6 shows the phase velocity for glass wool of density 30 kg/m^3, measured with 10-cm-thickness of glass wool. Measurement was also done for thicknesses 20 and 30 cm, and no dependence of phase velocity on thickness was found.

IV. CONCLUSION

The attenuation coefficients and phase velocities for glass wool of densities 14 and 30 kg/m^3 were measured. Attenuation is highest for waves propagating perpendicular to the fibers, the *X*-direction. The acoustic properties perpendicular to the *X*-direction are approximately isotropic in the frequency range 0–3000 Hz. For density 14 kg/m^3 the attenuation coefficient at 1000 Hz is 75 dB/m in the *X*-direction and 57 dB/m perpendicular to this direction. The elastic moduli was 2.0 and 120 kPa in the two respective directions.

[1] R. Kirby and A. Cummings, ''Prediction of the bulk properties of fibrous materials at low frequencies,'' Appl. Acoust. **56**, 101–125 (1999).
[2] M. E. Delany and E. N. Bazley, ''Acoustical properties of fibrous materials,'' Appl. Acoust. **3**, 105–116 (1970).
[3] V. Tarnow, ''Measurement of sound propagation in glass wool,'' J. Acoust. Soc. Am. **97**, 2272–2281 (1995).
[4] R. F. Lambert, ''Low-frequency acoustic behavior of highly porous, layered, flexible, fine fiber materials,'' J. Acoust. Soc. Am. **97**, 818–821 (1995).
[5] N. Voronina, ''Acoustic properties of fibrous materials,'' Appl. Acoust. **42**, 165–174 (1994).

[6] A. Bokor, ''Attenuation of sound in lined ducts,'' J. Sound Vib. **10**(3), 390–403 (1969).
[7] J. F. Allard, R. Bourdier, and A. L'Esperance, ''Anisotropy effect in glass wool on normal impedance in oblique incidence,'' J. Sound Vib. **114**(2), 233–238 (1987).
[8] V. Tarnow, ''Fiber movements and sound attenuation in glass wool,'' J. Acoust. Soc. Am. **105**, 234–240 (1999).
[9] K. Attenborough, ''Acoustical characteristics of rigid fibrous absorbents and granular materials,'' J. Acoust. Soc. Am. **73**, 785–799 (1983).
[10] J. F. Allard, *Propagation of Sound in Porous Media* (Elsevier Applied-Science, London, 1993).
[11] D. L. Johnson, J. Koplik, and R. Dashen, ''Theory of dynamic permeability and tortuosity in fluid-saturated porous media,'' J. Fluid Mech. **176**, 379–402 (1987).
[12] D. K. Wilson, ''Relaxation-matched modeling of propagation through porous media, including fractal pore structure,'' J. Acoust. Soc. Am. **94**, 1136–1145 (1993).
[13] V. Tarnow, ''Compressibility of air in fibrous materials,'' J. Acoust. Soc. Am. **99**, 3010–3017 (1996).

Optimum near-field performance of microphone arrays subject to a far-field beampattern constraint

James G. Ryan[a)]
Acoustics and Signal Processing, Institute for Microstructural Sciences, National Research Council Canada, Ottawa, Ontario K1A 0R6, Canada

Rafik A. Goubran
Systems and Computer Engineering, Carleton University, Ottawa, Ontario K1S 5B6, Canada

(Received 8 September 1999; revised 15 August 2000; accepted 15 August 2000)

This article addresses the problem of maximizing the near-field gain of a microphone array subject to a constraint on the far-field beampattern. The problem arises when acquiring speech from a near-field talker in the presence of a strong source of interference located farther from the array. When the angles of incidence from the near-field target and the far-field interference are identical, enforcing a null constraint in the interference direction reduces array gain and robustness. This article shows how to mitigate this effect by selection of the constrained response level. A suitable selection is to force the beampattern in the interference direction to be proportional to the unconstrained beampattern. The proportionality constant can then be used to trade off interference reduction and array gain. Specific numerical examples are provided. © *2000 Acoustical Society of America.* [S0001-4966(00)04711-1]

PACS numbers: 43.60.Gk, 43.38.Hz [JCB]

I. INTRODUCTION

This article addresses the problem of maximizing the near-field gain of a microphone array subject to a constraint on the far-field response pattern. It specifically deals with situations where the near-field signal source and the far-field interference are located along the same angle of incidence with respect to the array axis and both are immersed in a spherically isotropic noise field. This arises in speech pickup in rooms, for example, when a strong wall reflection or noise source originates from a location behind the near-field talker. The direct wave from both target and interference arrive at the array with the same angle of incidence; however, the interfering wavefront possesses a larger radius of curvature due to the increased distance between the source and the array.

Currently available methods simplify the analysis and design of microphone arrays by assuming that all acoustic sources are located far from the array. In such cases wavefront curvature observed at the array position is very small and all waves impinging upon the array are assumed to be planar.

There are circumstances, however, where the signal source is located very close to the array and the assumption of plane wave propagation is not valid.[1] This situation arises, for instance, in the application of microphone arrays to mobile telephony[2,3] and conference telephony in small rooms.[4,5] In these cases, the talker is located in the array's near field and wavefront curvature can be detected within the array's aperture.

Beamformer synthesis methods that explicitly include target wavefront curvature have recently been investigated. Current approaches provide a beamformer either with a desired spatial response pattern[6–8] or with an optimal response based on a suitable figure of merit.[2,9]

The problem of imposing a beampattern constraint on an optimized array has been addressed only in the case of far-field targets. Optimization approaches considered include maximizing array gain,[10,11] maximizing signal-to-noise ratio (SNR),[12] minimizing the sensor weight-vector norm,[13,14] and minimizing the beampattern disturbance caused by the added constraint.[15,16] Since all of these problem formulations involve the optimization of a quadratic performance index subject to a linear constraint, the solutions are related under the appropriate conditions.[13,17,16]

Steyskal[18] has addressed a mixed near-field/far-field problem which arises in the design of antenna systems. The objective was to minimize the effects of a near-field scatterer on the radiated beampattern of a phased array antenna. Using an extension of the method in Ref. 16 it was shown that a near-field response null can be imposed with minimal impact on the far-field beampattern.

The idea proposed in this work is essentially the opposite of Steyskal's, that is, to impose a far-field null while maintaining the response to a near-field target. The main difference is that the optimization goal is maximizing the near-field array gain rather than minimizing the pattern disturbance. Consequently, the objective is similar to that in Refs. 10 and 11; however, the derivation uses the results from linearly constrained minimum variance (LCMV) beamforming[19] and incorporates the concept of generalized angle found in Ref. 20. Furthermore, the result is generalized to include an arbitrary response level in the interference direction rather than a null. As is shown, proper selection of the constrained response level can reduce the impact of the constraint on the overall array performance.

This article is organized as follows. Concepts of conven-

[a)]Present address: Communications Research Center, 3701 Carling Ave., Ottawa, Ontario K2H 8S2, Canada.

tional array optimization methods are first outlined. Both unconstrained as well as linearly constrained optimization are reviewed. Next, the sensor weight vector and array gain are derived for the specific case of dual linear constraints: one representing the signal direction and another representing the interference or noise direction. The derivation generalizes that found in Refs. 10 and 11 by including an arbitrary response level at the interference location (as opposed to forcing a zero response). The generalized array-gain expression is then used to derive a noise-response target that permits a controlled trade-off of array gain. The results are illustrated using the specific example of a uniformly spaced linear array.

II. REVIEW OF UNCONSTRAINED AND LINEARLY CONSTRAINED ARRAY-GAIN MAXIMIZATION

Array gain G is defined as the ratio of output SNR over the SNR observed at a single sensor. For an M-element array, it is given by

$$G=\frac{\mathbf{W}^H\mathbf{R}_{ss}\mathbf{W}}{\mathbf{W}^H\mathbf{R}_{nn}\mathbf{W}}, \tag{1}$$

where $\mathbf{W}$ is the $M\times 1$ vector of complex sensor weights

$$\mathbf{W}^T=[W_0 \quad W_1 \quad \cdots \quad W_{M-1}], \tag{2}$$

and $\mathbf{R}_{ss}$ and $\mathbf{R}_{nn}$ are the $M\times M$ signal and noise correlation matrices defined as

$$\mathbf{R}_{ss}=E\{\mathbf{S}\cdot\mathbf{S}^H\} \tag{3}$$

and

$$\mathbf{R}_{nn}=E\{\mathbf{N}\cdot\mathbf{N}^H\}$$

with $\mathbf{S}$ being the signal propagation vector and $\mathbf{N}$ the noise propagation vector. The superscripts ^{T} and ^{H} denote matrix transpose and conjugate transpose, respectively.

The complex microphone weights defined in Eq. (2) apply to a specific frequency. Since speech signals are broadband, a practical implementation of these weights must be performed in the frequency domain [using a fast Fourier transform (FFT)] or by designing a finite impulse response (FIR) filter that provides the correct amplitude and phase at frequencies within the band of interest.

A. Unconstrained gain maximization

Provided that $\mathbf{R}_{nn}$ is nonsingular, the value of the quotient Eq. (1) is bounded by the minimum and maximum eigenvalues of the symmetric matrix $\mathbf{R}_{nn}^{-1}\mathbf{R}_{ss}$. The array gain is maximized by setting the weight vector $\mathbf{W}$ equal to the eigenvector corresponding to the maximum eigenvalue.[21] In the special case where $\mathbf{R}_{ss}$ is a dyad defined by the outer product

$$\mathbf{R}_{ss}=\mathbf{S}\mathbf{S}^H,$$

then the weight vector $\mathbf{W}_{\text{opt}}$ that maximizes G is given simply by[22]

$$\mathbf{W}_{\text{opt}}=\chi\mathbf{R}_{nn}^{-1}\mathbf{S}, \tag{4}$$

where χ is an arbitrary complex constant. Inclusion of the complex constant χ implies that the optimum weight vector is only unique to within a constant scale factor. Thus, the array gain is insensitive to a scaling of the weights. This added degree of freedom in selecting $\mathbf{W}_{\text{opt}}$ can be used to control the amplitude of the array output or to provide a flat frequency response for a broadband beamformer. The maximum array gain G_{opt} is

$$G_{\text{opt}}=\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{S}. \tag{5}$$

For certain noise correlation matrices, the unconstrained optimum weight solution Eq. (4) can lead to a beamformer that is intolerant of sensor noise and other types of random errors. Improved robustness can be obtained, however, by including a small diagonal term in the definition of $\mathbf{R}_{nn}$. That is, define

$$\mathbf{R}_{nn}=\mathbf{Q}+\delta^2\mathbf{I}, \tag{6}$$

where $\mathbf{Q}$ is the correlation matrix of the environmental noise, $\mathbf{I}$ is an $M\times M$ identity matrix, and δ^2 is a small positive constant used to control beamformer robustness. This modification allows simple, unconstrained optimization to be used to design a beamformer with an effective constraint on robustness.

This array optimization approach has been used to enhance the far-field performance of microphone arrays in noisy and reverberant environments.[23–25] Fixed, optimum arrays were designed based on an isotropic model of the interfering field and a plane wave model for the target. Results obtained with these optimized arrays demonstrate enhanced reception of speech signals in realistic situations. The extension of this concept to near-field targets is capable of providing a further increase in array gain.[9]

B. Linearly constrained gain maximization

To exert control over the array beampattern while optimizing performance, it is necessary to employ linear constraints. While various types of linear constraints are possible, this article focuses on the use of point constraints to constrain the response at a specific locations. Nonetheless, derivative and eigenvector constraints can also be used to exert broader control over the response shape.[19]

A single, linear constraint is expressed as

$$\mathbf{c}^H\mathbf{W}=g, \tag{7}$$

where $\mathbf{c}$ is an $M\times 1$ constraint vector and g is the desired response level. Point-constraint vectors correspond to the signal propagation vectors associated with the spatial constraint locations. For example, to constrain a linear array's far-field beampattern at an angle θ the following point constraint is required:

$$\mathbf{c}^T=[e^{jkx_0\cos\theta} \quad e^{jkx_1\cos\theta} \quad \cdots \quad e^{jkx_{M-1}\cos\theta}],$$

where x_m are the positions of the microphones and k is the wave number.

Multiple constraints can be enforced by defining a constraint matrix $\mathbf{C}$, whose L columns consist of individual constraint vectors $\mathbf{c}$, and a vector $\mathbf{g}$ of the corresponding responses g. The multiple constraint equation is

$$\mathbf{C}^H\mathbf{W}=\mathbf{g}.$$

The constrained optimization problem is then posed as

$$\text{maximize } G, \text{ subject to } \mathbf{C}^H\mathbf{W}=\mathbf{g}.$$

The method of Lagrange multipliers can be used to transform this constrained optimization problem into an unconstrained one. The problem is simplified by restating the gain maximization goal in an equivalent form. That is, maximizing array gain is equivalent to minimizing the noise response, as represented by the denominator of Eq. (1), with a constraint on the signal response, as represented by the numerator of Eq. (1). The signal response constraint can be handled easily as one of the columns in **C**. The optimization problem then becomes

$$\text{minimize } \mathbf{W}^H\mathbf{R}_{nn}\mathbf{W}+\boldsymbol{\lambda}^H(\mathbf{C}^H\mathbf{W}-\mathbf{g})+\boldsymbol{\lambda}^T(\mathbf{C}^H\mathbf{W}-\mathbf{g}),$$

where $\boldsymbol{\lambda}$ is the Lagrange multiplier vector. The weight vector that solves this problem is given by the well-known solution[19]

$$\mathbf{W}_{\text{opt}}=\mathbf{R}_{nn}^{-1}\mathbf{C}[\mathbf{C}^H\mathbf{R}_{nn}^{-1}\mathbf{C}]^{-1}\mathbf{g}. \tag{8}$$

Under equivalent conditions, the constrained optimum solution of Eq. (8) yields the same solution as the unconstrained solution Eq. (4). To see this, consider the case of a single point constraint at the signal location $\mathbf{C}=\mathbf{S}$. The constraint equation is

$$\mathbf{S}^H\mathbf{W}=g_s,$$

where g_s is the desired signal level. According to Eq. (8) the linearly constrained optimum weights are

$$\mathbf{W}_{\text{opt}}=\frac{\mathbf{R}_{nn}^{-1}\mathbf{S}}{\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{S}}g_s, \tag{9}$$

which is equal to the unconstrained solution given in Eq. (4) with

$$\chi=\frac{g_s}{\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{S}}$$

confirming that the two optimization approaches yield equivalent results.

III. NEAR-FIELD GAIN MAXIMIZATION SUBJECT TO A FAR-FIELD BEAMPATTERN CONSTRAINT

This article examines the impact of a single, linear constraint on the performance of an optimized array. The derivation presented herein allows the specification of an arbitrary response level. As discussed earlier, however, the optimum array weights are insensitive to scaling by a constant factor. Without further precautions, the solution to the constrained optimization problem would be to simply scale the entire unconstrained beampattern. Both the target and interference would be scaled by the same amount. To avoid this situation, this derivation includes two linear constraints. One constraint is used to control the beamformer response in the target direction **S** while the other is used to control the response in a known interference direction **N**. The only assumption made about **S** and **N** is that they are not identical (otherwise the matrix represented by $\mathbf{C}^H\mathbf{R}_{nn}^{-1}\mathbf{C}$ would not be invertible). In the following material, results will be demonstrated for a near field **S** and a far field **N**; however, the results apply equally well to **S** and **N** in the near field or far field.

A. Optimum weights

Define the $M\times 2$ constraint matrix as

$$\mathbf{C}=[\mathbf{S}\quad\mathbf{N}],$$

where **S** and **N** represent assumed signal and noise vectors, respectively. The 2×1 gain vector is

$$\mathbf{g}=\begin{bmatrix} g_s \\ g_n \end{bmatrix}.$$

The effect of this constraint system is to force the beamformer to have a response of g_s at a location corresponding to vector **S** and a response of g_n at a location corresponding to vector **N**. The value of g_s is chosen to provide a desired signal level at the array output and the value of g_n is chosen to suppress the interference. In principle, the values of both these parameters can be set arbitrarily; however, the choice of g_n has an impact on overall beamformer performance, as will be discussed later.

Substituting the constraints into the expression for the optimum weights Eq. (8) yields

$$\mathbf{W}_{\text{opt}}=[\mathbf{R}_{nn}^{-1}\mathbf{S}\quad\mathbf{R}_{nn}^{-1}\mathbf{N}]\begin{bmatrix}\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{S} & \mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{N}\\ \mathbf{N}^H\mathbf{R}_{nn}^{-1}\mathbf{S} & \mathbf{N}^H\mathbf{R}_{nn}^{-1}\mathbf{N}\end{bmatrix}^{-1}\begin{bmatrix} g_s \\ g_n \end{bmatrix}.$$

The inverse of the 2×2 matrix can be written as

$$\begin{bmatrix}\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{S} & \mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{N}\\ \mathbf{N}^H\mathbf{R}_{nn}^{-1}\mathbf{S} & \mathbf{N}^H\mathbf{R}_{nn}^{-1}\mathbf{N}\end{bmatrix}^{-1}$$

$$=\frac{1}{D}\cdot\begin{bmatrix}\mathbf{N}^H\mathbf{R}_{nn}^{-1}\mathbf{N} & -\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{N}\\ -\mathbf{N}^H\mathbf{R}_{nn}^{-1}\mathbf{S} & \mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{S}\end{bmatrix},$$

where D is the matrix determinant

$$D=[\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{S}][\mathbf{N}^H\mathbf{R}_{nn}^{-1}\mathbf{N}]-|\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{N}|^2.$$

Substituting the expression for the inverse matrix, multiplying through, and rearranging terms yields the following solution for the optimum weights

$$\mathbf{W}_{\text{opt}}=\frac{1}{1-(\cos\gamma)^2}\cdot[g_s\mathbf{W}_s-g_n\mathbf{W}_s\mathbf{S}^H\mathbf{W}_n+g_n\mathbf{W}_n-g_s\mathbf{W}_n\mathbf{N}^H\mathbf{W}_s]. \tag{10}$$

In this expression, $\mathbf{W}_s$ is the unconstrained optimum weight vector for a source located at **S** [see Eq. (9)],

$$\mathbf{W}_s=\frac{\mathbf{R}_{nn}^{-1}\mathbf{S}}{\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{S}},$$

$\mathbf{W}_n$ is the unconstrained optimum weight vector for a source located at **N**,

$$\mathbf{W}_n=\frac{\mathbf{R}_{nn}^{-1}\mathbf{N}}{\mathbf{N}^H\mathbf{R}_{nn}^{-1}\mathbf{N}},$$

and $(\cos\gamma)^2$ is the cosine squared of the generalized angle γ between the constraint vectors **S** and **N** defined as[20]

$$(\cos\gamma)^2=\frac{|\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{N}|^2}{\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{S}\cdot\mathbf{N}^H\mathbf{R}_{nn}^{-1}\mathbf{N}}. \tag{11}$$

Note that, due to the Schwarz inequality,

$$0\leqslant(\cos\gamma)^2\leqslant 1.$$

When $(\cos\gamma)^2=0$, the vectors **S** and **N** are orthogonal in the generalized sense. Likewise when $(\cos\gamma)^2=1$, the vectors are colinear. When $\mathbf{R}_{nn}$ is equal to the identity matrix **I**, the generalized angle reduces to a more familiar form.

Note the denominator term involving $(\cos\gamma)^2$. If the vectors **S** and **N** are orthogonal, then $(\cos\gamma)^2=0$ and the denominator term has no effect. As $(\cos\gamma)^2\rightarrow 1$, however, the denominator approaches 0 and the weight magnitudes increase drastically. Under this condition, the resulting array has very low robustness.

B. Optimum gain

By substituting the expression Eq. (10) for $\mathbf{W}_{\text{opt}}$ into Eq. (1), a specific expression can be derived for the expected gain of the beamformer. The details are tedious but straightforward and are not provided here. The result is the optimum array gain for the linearly constrained beamformer

$$G=G_u\cdot\frac{|g_s|^2(1-(\cos\gamma)^2)}{|g_s|^2+|g_n|^2\cdot(\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{S}/\mathbf{N}^H\mathbf{R}_{nn}^{-1}\mathbf{N})-[2\,\text{Re}\{g_sg_n^*(\mathbf{N}^H\mathbf{R}_{nn}^{-1}\mathbf{S})\}]/[\mathbf{N}^H\mathbf{R}_{nn}^{-1}\mathbf{N}]}, \tag{12}$$

where

$$G_u=\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{S}$$

is the unconstrained optimum gain obtained for a target direction of **S** [see Eq. (5)], that is, the gain of an array optimized for reception of **S** subject to a single gain constraint in the look direction and no second constraint in the direction of **N**. While it is difficult to determine from inspection of this expression, constrained array gain is always less than or equal to the unconstrained gain.[10] This result applies equally to the white noise gain, a common measure of robustness, since no assumption was made about $\mathbf{R}_{nn}$. Consequently, imposing a linear constraint on an optimized beamformer will degrade both the array gain as well as the beamformer's robustness, although to different degrees.

For the specific example of a signal constraint $\mathbf{S}^H\mathbf{W}_s=g_s=1$ and a null constraint $\mathbf{N}^H\mathbf{W}_n=g_n=0$, $\mathbf{W}_{\text{opt}}$ simplifies to

$$\mathbf{W}_{\text{opt}}=\frac{1}{1-(\cos\gamma)^2}\cdot[\mathbf{W}_s-\mathbf{W}_n\mathbf{N}^H\mathbf{W}_s]. \tag{13}$$

The quantity $\mathbf{N}^H\mathbf{W}_s$ is the response of a beamformer optimized for reception of **S** to the direction **N**. This is precisely the amount of component **N** that must be cancelled in the response of the $\mathbf{W}_s$ beamformer. Thus, the linearly constrained optimum solution is composed of the sum of an optimum beamformer $\mathbf{W}_s$ and a nulling beamformer $-\mathbf{W}_n\mathbf{N}^H\mathbf{W}_s$, in agreement with previous results.[11] The array gain for this example simplifies to

$$G=G_u\cdot(1-(\cos\gamma)^2).$$

Thus, in the case of a single null constraint, the resulting array gain is smaller than the gain of an equivalent unconstrained array unless **S** and **N** are orthogonal $[(\cos\gamma)^2=0]$.

For many of the examples encountered in the literature, the vectors **S** and **N** are orthogonal or nearly orthogonal and the impact on array gain is small.[10,16] For a near-field **S** and far-field **N** with the same angle of incidence, however, the vectors are rarely orthogonal and significant gain penalties can result, depending on the chosen value of g_n.

Fortunately, it is possible to select the value of g_n such that the impact on array gain, or robustness, is minimized. Using this approach, the value of g_n can be used to trade-off array gain for the suppression of discrete interferences. This point is discussed in the following section.

C. Selecting constraint level to limit gain loss

For a given pair of constraint vectors **S** and **N**, it is possible to select the gain parameter g_n such that the array gain is maximized. Referring to Eq. (12), it is seen that only the denominator is a function of g_n. Consequently, this expression can be maximized by choosing the value of g_n that minimizes the denominator. This value can be found by taking the derivative of the denominator expression with respect to g_n and setting the result equal to 0. Solving for g_n yields

$$g_n|_{\text{opt}}=g_s\frac{\mathbf{N}^H\mathbf{R}_{nn}^{-1}\mathbf{S}}{\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{S}}=g_s\mathbf{N}^H\mathbf{W}_s.$$

As seen, $g_n|_{\text{opt}}$ is equal to the response of an array optimized for reception of **S** to a source corresponding to **N**. It is identical to the response which would have been obtained in the absence of a constraint on **N**. Thus, the choice of g_n that minimizes impact on the array gain is the response of the unconstrained beamformer to the null direction, **N**, the same response that would be obtained in the absence of the constraint on **N**. All other values of g_n will result in a lower *G*.

Unfortunately, the optimal g_n provides no added attenuation of **N**; however, the result suggests a compromise value of g_n. By selecting g_n as a fraction of the maximizing value, the response to **N** can be reduced while controlling the gain loss to any desired level. That is, choose

$$g_n=\alpha g_s\frac{\mathbf{N}^H\mathbf{R}_{nn}^{-1}\mathbf{S}}{\mathbf{S}^H\mathbf{R}_{nn}^{-1}\mathbf{S}}, \tag{14}$$

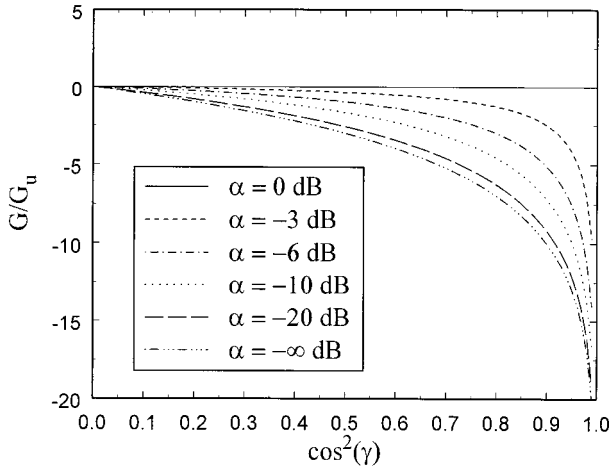


FIG. 1. Gain loss due to linear constraint imposition as a function of $(\cos\gamma)^2$ for several values of α.

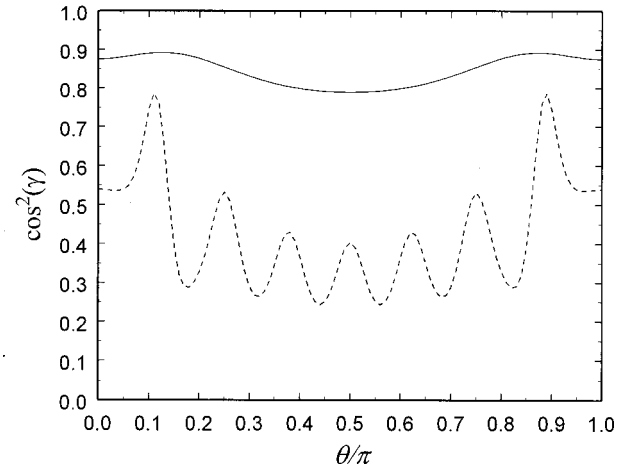


FIG. 2. Variation in $(\cos\gamma)^2$ between near-field point at a distance L and far-field as a function of angle of incidence with respect to array axis θ. Solid line assumes uncorrelated noise (i.e., $\mathbf{R}_{nn}=\mathbf{I}$) whereas dashed line assumes isotropic noise.

where α is a constant which determines the additional interference suppression. Typically, it would be selected somewhere between 0 and 1. Using Eq. (14) in Eq. (12) yields

$$G=G_u\frac{(1-(\cos\gamma)^2)}{-[1-(2\alpha-\alpha^2)(\cos\gamma)^2]}.$$

To provide some indication of the gain loss due to constraints, the ratio G/G_u is calculated for various values of α. The results are plotted in Fig. 1 as a function of $(\cos\gamma)^2$. The data in this figure show that substantial gain loss occurs as $(\cos\gamma)^2\to 1$. The loss in gain decreases as $(\cos\gamma)^2\to 0$ and is zero for constraint vectors that are orthogonal $(\cos\gamma)^2=0$. Furthermore, for a given value of $(\cos\gamma)^2$ greater loss of array gain occurs when enforcing constraints for small values of α. For example, limiting gain loss to less than 3 dB requires that $(\cos\gamma)^2\leqslant 0.5$ when $\alpha=-30$ dB. If the constraint is loosened to $\alpha=-3$ dB, however, it is only necessary that $(\cos\gamma)^2\leqslant 0.9$. The implications for suitable array focal and constraint points is discussed next with reference to a specific example.

The primary interest in this article is to examine the application of the constrained optimization technique to near-field sources. More specifically, when the signal propagation vector **S** corresponds to a near-field point source and the noise propagation vector **N** corresponds to a plane-wave incident from the same angle as **S**. The feasibility of a far-field response constraint is illustrated by examining the variation of $(\cos\gamma)^2$ for various **S** within the near field of a typical array. The specific example considered is that of an eight-element uniformly spaced linear array focused at a distance equal to the array length. The variation in $(\cos\gamma)^2$ is plotted in Fig. 2 as a function of angle of incidence θ with respect to the array axis. The two curves illustrate the effect of different $\mathbf{R}_{nn}$ on the generalized angle. The solid curve was obtained assuming uncorrelated noise between all sensors, that is, $\mathbf{R}_{nn}=\mathbf{I}$. The dashed curve assumes $\mathbf{R}_{nn}=\mathbf{Q}$ where **Q** is the noise correlation matrix corresponding to a spherically isotropic noise field.[26]

As seen in Fig. 2, the vectors **S** and **N** are not orthogonal in the generalized sense for any angle of incidence. This implies that there will always be a gain reduction when performing a constrained optimization. Furthermore, the nature of $\mathbf{R}_{nn}$ has a significant impact on $(\cos\gamma)^2$. For the array geometry and source positions considered here the use of $\mathbf{R}_{nn}=\mathbf{Q}$ orthogonalizes the vectors **S** and **N** compared to $\mathbf{R}_{nn}=\mathbf{I}$. Referring to Fig. 1, this will moderate the loss in array gain due to the constraint.

It should be noted that a realistic beamformer design will not use either of these noise correlation matrices in isolation. An optimum design produced for $\mathbf{R}_{nn}=\mathbf{I}$ produces an array with little directional gain whereas an array optimized using $\mathbf{R}_{nn}=\mathbf{Q}$ is not robust. A more practical choice for $\mathbf{R}_{nn}$ will include components of both **Q** and **I**, as discussed in the next section. In such cases, the curves in Fig. 2 represent upper and lower bounds for $(\cos\gamma)^2$. The actual angle between vectors will lie somewhere within this range depending upon the relative amounts of the **Q** and **I** components.

IV. NUMERICAL EXAMPLES

The preceding optimizations can be compared using the specific example of an eight-microphone linear array with $\lambda/4$ sensor spacing. The array is focused at a distance equal to the array length, L, in the broadside direction ($\theta=\pi/2$), as illustrated in Fig. 3. The noise correlation matrix is defined as in Eq. (6) with **Q** being the correlation matrix corresponding to a spherically isotropic field[26] and $\delta^2=0.0103$. This value of δ^2 results in a white noise gain of 0 dB in the unconstrained optimized beamformer [designed according to Eq. (4)]. Since $\mathbf{R}_{nn}$ contains both a spherically isotropic portion as well as a spatially incoherent portion, the exact $(\cos\gamma)^2$ for **S** and **N** will be somewhere in between the extremes depicted in Fig. 2. For an angle of incidence of $\pi/2$, $(\cos\gamma)^2=0.4$ when $\mathbf{R}_{nn}=\mathbf{Q}$ and $(\cos\gamma)^2=0.8$ when $\mathbf{R}_{nn}=\mathbf{I}$.

Three beamformers are considered: an unconstrained optimum, a null-constrained optimum, and a constrained op-

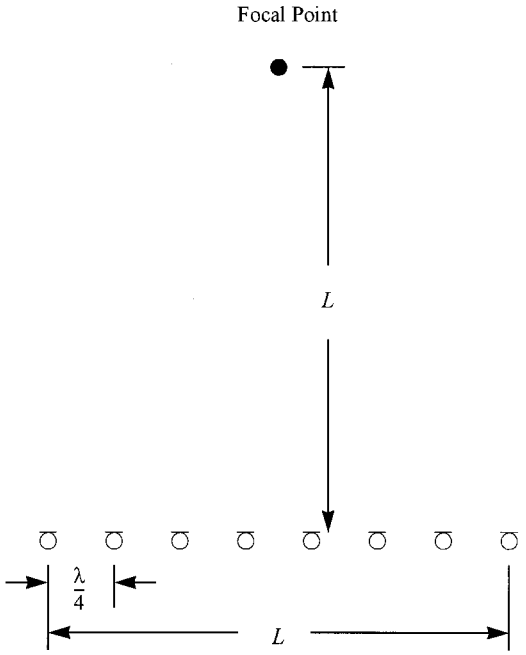


FIG. 3. Illustration of an eight-element, equispaced linear array and near-field focal-point position.

timum with reduced interference suppression.

A. Unconstrained optimum

The unconstrained optimum beamformer is designed according to Eq. (4) with no auxiliary constraints on the response pattern. Near-field and far-field beampatterns for the optimized array are shown in Fig. 4. The near-field beampattern is evaluated at a distance r equal to the focal distance L. It can be seen that the beamformer exhibits a pronounced peak in the response at the correct focal point. It is also seen that the far-field response pattern (evaluated at $r=\infty$) is less directional but still responds strongly in the focal direction ($\pi/2$). The array gain for this case is 6.6 dB.

B. Null constrained

The next example is a beamformer optimized subject to a null constraint in the far-field direction as described by Eq. (13). Beampatterns for the null-constrained optimum design are shown in Fig. 5. These should be compared with similar beampatterns obtained for the unconstrained optimum array shown in Fig. 4. The null-constrained design provides the desired response at the focal point; however, the near-field beampattern possesses somewhat higher sidelobe levels. In keeping with the null constraint, the far-field beampattern displays a sharp null in the response in the broadside direction.

The cost of this added null in the far-field pattern is evident in the array gain which is reduced from 6.6 dB, for the unconstrained case, down to 2.3 dB a loss of 4.3 dB. This is within the range of 2.2 to 7.2 dB predicted by the $\alpha=-\infty$ dB curve in Fig. 1. Furthermore, the null-constrained beamformer also possesses reduced robustness. The white noise gain is reduced by 10.4 dB compared to the unconstrained example.

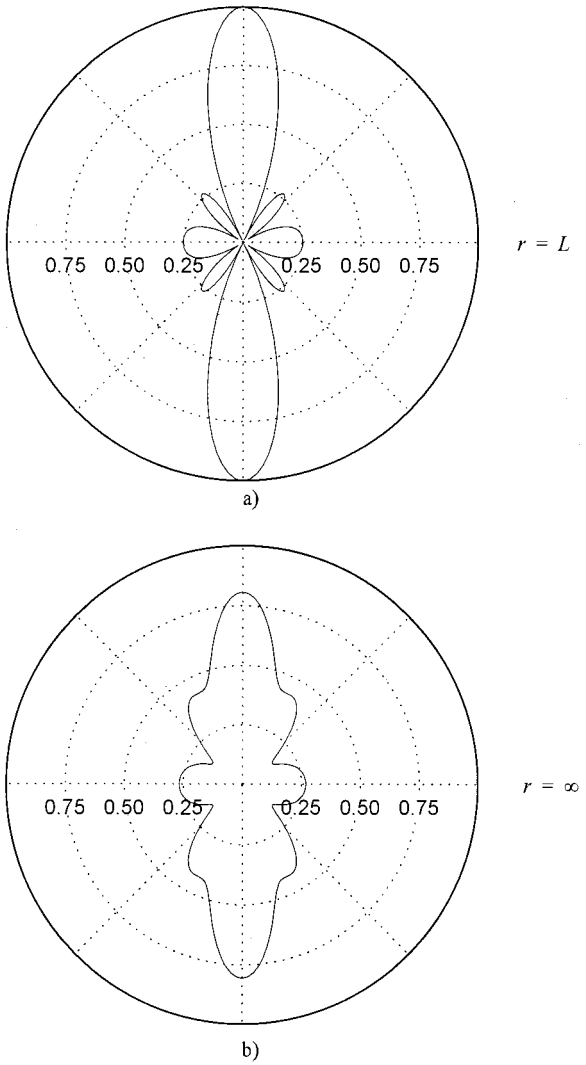


FIG. 4. Near-field and far-field beampatterns for an eight-element uniformly spaced linear array with λ/4 sensor spacing. The array is focused at $(r_f, \theta_f)=(L, \pi/2)$ and the beampattern is evaluated at distances $r=L$ (a) and $r=\infty$ (b).

C. Constrained to limit gain reduction

Beampatterns for the same array but designed using the constraint in Eq. (14) are shown in Fig. 6. For this example, the constraint parameter α is set to -6 dB. Incorporating the soft constraint lessens the impact on the near-field beampattern's sidelobe levels compared to the null constraint. Furthermore, the far-field response has been reduced by the -6 dB compared to the beampattern in Fig. 4, as expected.

The benefit of the soft constraint is that, in this case, the array gain is reduced by only 1.6 dB. This is in keeping with the predicted loss of between 0.7 and 3.1 dB. Furthermore, the white noise gain reduction is limited to 6.4 dB, resulting in a more-robust design.

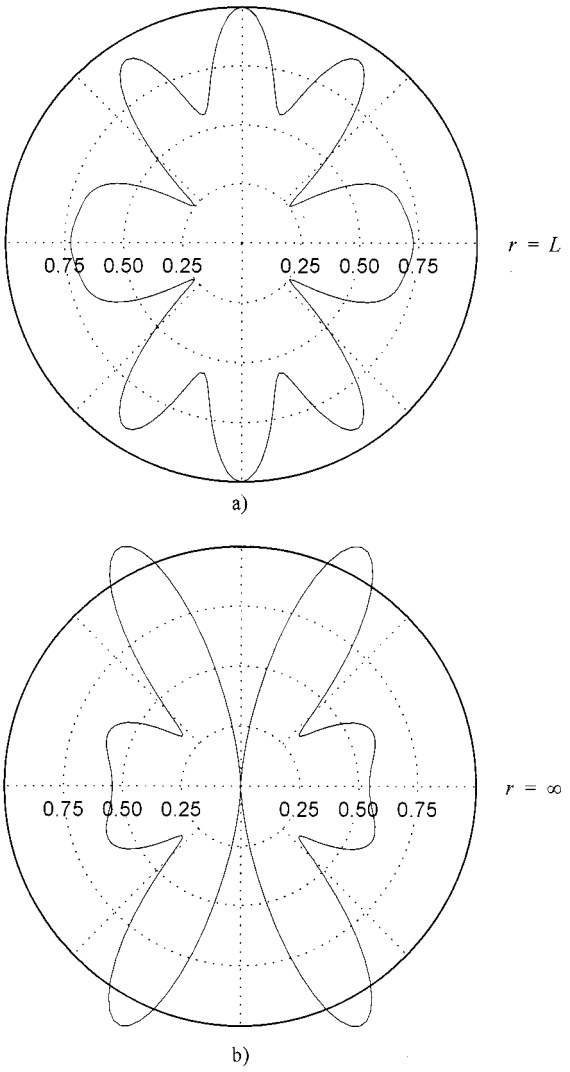


FIG. 5. Beampatterns for the same array as in Fig. 4 but using linearly constrained optimal weights. Array gain is optimized while forcing a null in the broadside direction. The beampattern is evaluated at distances $r=L$ (a) and $r=\infty$ (b).

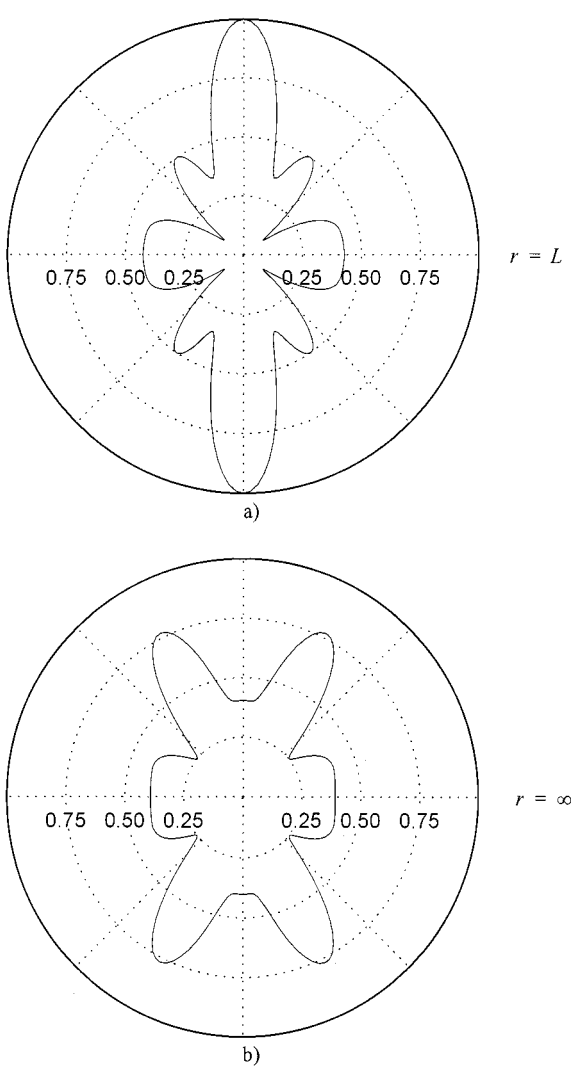


FIG. 6. Soft-constrained optimum beampatterns for the array in Fig. 4. Array gain is optimized while suppressing the response in the broadside direction by an additional 6 dB compared to that shown in Fig. 4. The beampattern is evaluated at distances $r=L$ (a) and $r=\infty$ (b).

V. CONCLUSIONS

This article addresses the problem of maximizing the near-field gain of a microphone array subject to a constraint on the far-field beampattern. It specifically addresses situations where the near-field signal source and the far-field interference are located along the same angle of incidence with respect to the array axis and both are immersed in a spherically isotropic field. This arises in speech pickup in rooms, for example, when a strong wall reflection or noise source originates from a location behind the near-field talker.

When the angles of incidence from the near-field target and the far-field interference are identical, enforcing a null constraint in the interference direction reduces array gain and robustness. This article shows that this effect can be mitigated by proper selection of the beampattern constraint in the interference direction. By constraining the beampattern in the interference direction to be proportional to that of the unconstrained beampattern a minimum loss in array gain and robustness are achieved. The proportionality constant can then be used to trade off interference reduction and array gain.

Using the example of an eight-element linear array focused in the broadside direction, the beampatterns and array gain of unconstrained and linearly constrained optimum beamformers are compared. Forcing a null in the interference direction results in a gain reduction of 4.3 dB. By constraining the interference to a level 6 dB below the unconstrained beamformer, the loss in array gain is limited to 1.6 dB. The new constraint also results in a more robust beamformer as indicated by the white noise gain.

ACKNOWLEDGMENTS

The authors would like to thank Dr. M. R. Stinson of the National Research Council of Canada for helpful comments on an earlier version of this manuscript.

[1] J. G. Ryan, ''Criterion for the minimum source distance at which plane-wave beamforming can be applied,'' J. Acoust. Soc. Am. **104**, 595–598 (1998).

[2] M. M. Goulding and J. S. Bird, ''Speech enhancement for mobile telephony,'' IEEE Trans. Veh. Technol. **39**, 316–326 (1990).
[3] S. Nordholm, I. Claesson, and B. Bengtsson, ''Adaptive Array Noise Suppression of Handsfree Speaker Input in Cars,'' IEEE Trans. Veh. Technol. **42**, 514–518 (1993).
[4] F. Khalil, J. P. Jullien, and A. Gilloire, ''Microphone Array for Sound Pickup in Teleconference Systems,'' J. Audio Eng. Soc. **42**, 691–700 (1994).
[5] F. Pirz, ''Design of a Wideband, Constant Beamwidth, Array Microphone for Use in the Near Field,'' Bell Syst. Tech. J. **58**, 1839–1850 (1979).
[6] R. A. Kennedy, T. Abhayapala, and D. B. Ward, ''Broadband nearfield beamforming using a radial beampattern transformation,'' IEEE Trans. Signal Process. **46**, 2147–2156 (1998).
[7] D. B. Ward and G. W. Elko, ''Mixed near-field/far-field beamforming: A new technique for speech acquisition in a reverberant environment,'' *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics 1997*, October 1997.
[8] S. Nordholm, V. Rehbock, K. L. Teo, and S. Nordebo, ''Chebyshev optimization for the design of broadband beamformers in the near field,'' IEEE Trans. Circuits Syst. II: Analog Digital Signal Process. **45**, 141–143 (1998).
[9] J. G. Ryan and R. A. Goubran, ''Array optimization applied in the near field of a microphone array,'' IEEE Trans. Speech Audio Process **8**, 173–176 (2000).
[10] C. Drane and J. McIlvenna, ''Gain Maximization and Controlled Null Placement Simultaneously Achieved in Aerial Array Patterns,'' Radio Electron. Eng. **39**, 49–57 (1970).
[11] A. Fenwick, ''An array processor subject to a constraint,'' J. Acoust. Soc. Am. **49**, 1108–1110 (1971).
[12] P. C. K. Kwok and P. S. Brandon, ''Maximization of signal/noise ratio in antenna arrays subject to constraints,'' Proc. Inst. Electr. Eng. **126**, 1241–1244 (1979).
[13] T. S. Ng, J. Y. Chong Cheah, and F. J. Paoloni, ''Optimization with controlled null placement in antenna array pattern synthesis,'' IEEE Trans. Antennas Propag. **33**, 215–217 (1985).
[14] P. J. Kootsookos, D. B. Ward, and R. C. Williamson, ''Imposing pattern nulls on broadband array responses,'' J. Acoust. Soc. Am. **105**, 3390–3398 (1999).
[15] J. T. Mayhan and L. J. Ricardi, ''Physical limitations on interference reduction by antenna pattern shaping,'' IEEE Trans. Antennas Propag. **23**, 639–646 (1975).
[16] H. Steyskal, ''Synthesis of antenna patterns with prescribed nulls,'' IEEE Trans. Antennas Propag. **30**, 273–279 (1982).
[17] T. S. Ng, ''Equivalence of null constrained optimization problems in antenna array pattern synthesis,'' J. Electr. Electron. Eng., Aust. **7**, 302–305 (1987).
[18] H. Steyskal, ''Synthesis of antenna patterns with imposed near-field nulls,'' Electron. Lett. **30**, 2000–2001 (1994).
[19] B. D. Van Veen and K. M. Buckley, ''Beamforming: a versatile approach to spatial filtering,'' IEEE Signal Process. Mag. , 4–24 (1988).
[20] H. Cox, ''Resolving power and sensitivity to mismatch of optimum array procesors,'' J. Acoust. Soc. Am. **54**, 771–785 (1973).
[21] R. A. Monzingo and T. A. Miller, *Introduction to Adaptive Arrays* (Wiley, New York, 1980).
[22] D. K. Cheng and F. I. Tseng, ''Gain optimization for arbitrary antenna arrays,'' IEEE Trans. Antennas Propag. **13**, 973–974 (1965).
[23] J. M. Kates, ''Superdirective arrays for hearing aids,'' J. Acoust. Soc. Am. **94**, 1930–1933 (1993).
[24] R. W. Stadler and W. M. Rabinowitz, ''On the Potential of fixed arrays for hearing aids,'' J. Acoust. Soc. Am. **94**, 1332–1342 (1993).
[25] W. Soede, F. A. Bilsen, and A. J. Berkhout, ''Development of a directional hearing instrument based on array technology,'' J. Acoust. Soc. Am. **94**, 785–798 (1993).
[26] C. T. Morrow, ''Point-to-point correlation of sound pressures in reverberation chambers,'' J. Sound Vib. **16**(1), 29–42 (1971).

Adaptive beamforming at very low frequencies in spatially coherent, cluttered noise environments with low signal-to-noise ratio and finite-averaging times

Albert H. Nuttall
Naval Undersea Warfare Center, Newport, Rhode Island 02841

James H. Wilson
Neptune Sciences, Inc., 3834 Vista Azul, San Clemente, California 92672

(Received 15 September 1998; revised 15 August 2000; accepted 15 August 2000)

Realistic simulations with spatially coherent noise have been run in order to compare the performance of adaptive beamforming (ABF), inverse beamforming (IBF), and conventional beamforming (CBF) for the case of finite-averaging times, where the actual spatial coherence of the acoustic field, or covariance matrix, is not known *a priori*, but must be estimated. These estimation errors cause large errors in the ABF estimate of the directionality of the acoustic field, partly because ABF is a highly nonlinear algorithm. In addition, it is shown that ABF is fundamentally limited in its suppression capability at very low frequency (VLF), based on the sidelobe level of the conventional beampattern in the direction of the noise interferer [G. L. Mohnkern, ''Effects of Errors and Limitations on Interference Suppression,'' *NOSC Technical Document 1478*, Naval Ocean Systems Center (1989)]. The simulations include a low-level plane wave signal of interest, a stronger noise plane wave interferer, and spatially random background noise. Both IBF and ABF performed significantly better than CBF, and IBF's performance was slightly better than ABF's performance. The performances of IBF and the ABF algorithm, the minimum variance distortionless response (MVDR) [A. H. Nuttall and D. W. Hyde, ''Unified Approach to Optimum and Suboptimum Processing for Arrays,'' *USL Report Number 992*, Naval Underwater Systems Center, New London, CT (22 April 1969)] were recently compared independently [J. S. D. Solomon, A. J. Knight, and M. V. Greening, ''Sonar Array Signal Processing for Sparse Linear Arrays,'' Defense Science and Technology Organization (DSTO) Technical Report (June 1999)] using measured data, with the result that IBF outperformed MVDR. This result is significant because MVDR requires orders of magnitude more processing power than IBF or CBF. [S0001-4966(00)04811-6]

PACS numbers: 43.60.Lq, 43.30.Wi, 43.30.Nb [JCB]

I. BACKGROUND

There has been a great deal of effort expended over the years since the pioneering works of Capon (1969) and Nuttall and Hyde (1969) in developing the minimum variance distortionless response (MVDR) adaptive beamforming (ABF) algorithm as the ''optimum'' beamformer for realistic noise environments where surface ship clutter makes the noise field spatially coherent. Many articles [see Burdic (1991) and Mohnkern (1989) for a representative list of references] have addressed MVDR, to the point where it is now implemented in U.S. Navy surveillance and tactical passive sonar systems. MVDR effectively maximizes the beam output signal-to-noise ratio (SNR) by means of minimizing the total output power while maintaining a constant gain in a specified beam direction. The intent of MVDR is to effectively suppress clutter noise due to multiple surface ships, by adaptively changing the array beampattern so that all surface ship sources are ''notched.''

MVDR initially assumes perfect knowledge of the noise spatial and temporal coherence, or covariance matrix. However, the performance of MVDR is degraded in practice when there are errors in the estimate of the noise spatial coherence, which always occurs when finite-averaging times are used in passive sonar systems with low SNR. Furthermore, MVDR is a highly nonlinear procedure, with the result that small errors in estimation of the true noise covariance matrix can cause large errors in MVDR's output; these errors can thwart attempts to effectively suppress large noise clutter while attempting to detect low SNR signals emanating from targets of interest.

The objective of this article is to perform a limited sensitivity analysis of the performance of a particular MVDR algorithm when relatively short, finite-averaging times are used to estimate the noise covariance matrix. A simulation model, described in the next section, was developed to address the case of realistic (i.e., spatially coherent) noise clutter from shipping. The simulation approach was adopted because a totally theoretical approach could not be formulated and analyzed, at least one that addresses all the MVDR system details and the complex spatial acoustic field. Theoretical analyses based on the assumptions of isotropic, statistically independent noise and perfect estimates of the noise covariance matrix produce results that can be very misleading. The results of theoretical analyses based on unrealistic noise assumptions do not represent the real ocean environment, and large expenditures implementing MVDR in pas-

sive sonar systems should be avoided until the beamforming performance of MVDR is fully assessed for low SNR, finite-averaging times, and at very low frequencies (VLF).

In this article, the simulator results will show that MVDR's performance is significantly degraded when short averaging times, typical of tactical passive sonars (10- to 20-s update intervals) are utilized. As the averaging time is increased, the estimate of the noise covariance matrix is expected to improve. However, practical ABF performance suffers from the fact that shipping noise sources often do not ''stay in the same beam'' long enough to provide an accurate, time-stationary estimate of the adaptive weights required to effectively suppress clutter sources. The conclusion of this article is that MVDR is the optimum beamformer only for the case of long-averaging times (e.g., surveillance arrays) and for noise sources at long ranges that ''stay in one beam'' long enough to allow an accurate, time-stationary estimate of the covariance matrix to be made.

For shallow water, where nearby shipping noise sources dominate and have relatively high bearing rates, or for deep-water sites, where nearby shipping dominates (e.g., Sea of Japan, Mediterranean Sea, Straits of Sicily, etc.), the simulator results in this article predict that MVDR may not perform well in practice in suppressing clutter sources or maximizing the output SNR in an ''optimum'' way. In practice, either the tactical passive sonar will be allowed only short-averaging times, or the dominant shipping noise sources will have too high a bearing rate to allow longer averaging times. In either case, the errors in the estimate of the noise covariance matrix can produce large errors in the MVDR clutter suppression output, mainly because MVDR is a highly nonlinear beamforming algorithm. A recent side-by-side performance comparison (Solomon *et al.*, 1999) using measured data showed that IBF outperformed ABF/MVDR and supports the independent results of the simulations presented in this article.

In shallow water at VLF, Wilson and Veenhuis (1997) have shown that plane wave beamformers are significantly degraded, especially near endfire and at the design frequency of a horizontal array, by modal propagation of the signal energy. In fact, the performance of all plane wave beamformers, including CBF, Fourier integral method (FIM), and MVDR, are impacted by modal propagation, but it is not yet known which ones are most affected. For shallow water environments and array orientations where plane waves are not good approximations to the arrival structure, the authors suggest the use of matched field processing (MFP) with flat weighting. At VLF, where the CBF beamwidths are rather wide, the use of flat element weighting can be instrumental to the success of MFP, due to its narrower mainlobe.

II. ABF PERFORMANCE AT VLF

As submarine emissions have become quieter overall, the higher frequency emissions have been almost eliminated. This has made the VLF spectrum (<100 Hz) an especially important area of interest for modern passive sonars. Without intending to prove it, Mohnkern (1989) showed that ABF's performance is fundamentally limited at VLF by the array's conventional beam pattern. Mohnkern used the MVDR algorithm for his ABF calculations. Expense, practical limitations of towed array operations, and shallow water are just three reasons why short aperture arrays must be used for the foreseeable future. At VLF, the conventional beamwidths become rather wide and, for many realistic cases, the signal of interest and the interferer are within the same conventional beam. Mohnkern (1989) derived an expression for the maximum tolerable conventional sidelobe response, g_q, in the direction of the interferer as a function of interferer level, s, and theoretical suppression factor, $1+q$ (using Mohnkern's notation). Mohnkern (1989) had a typographical error in his equation (8); the correct expression is

$$g_q = \frac{sMq+q}{(1+q)Ms}$$

for an M-element, equally-spaced line array. Mohnkern's (1989) Figure 1 is essentially correct at high interference to noise ratios, and shows that it is theoretically impossible to suppress the interferer's level below 3 dB above the background noise level at VLF when the interferer and signal of interest are within the same conventional beamwidth. The simulation results in this article are consistent with this theoretical limitation of MVDR at VLF. Errors in the estimation of the covariance matrix due to short, finite-averaging times further degrade MVDR performance at VLF, to the point where the orders of magnitude increase in processing power (and hardware) required by MVDR is not worth the return of slightly worse performance compared to other beamforming algorithms, such as the FIM, will be described and evaluated in the next section.

The major result of the realistic simulations in this article is that the nonadaptive FIM beamformer achieved equal or better performance to the MVDR element-based (EB) algorithms, even though MVDR requires an order of magnitude more processing power than FIM. Other authors (Vural, 1974; Kneipfer, 1997) have developed the alternative procedure called beam-based (BB) MVDR, which significantly reduces the computational load, but cannot improve the signal detectability above EB MVDR. Our direction instead is to develop an adaptive FIM algorithm related to the CLEAN (Thompson *et al.*, 1994) algorithm, which not only requires a minimal computational load, but also significantly improves the signal detectability performance under the practical conditions of low input SNR and short observation times.

The fact that the FIM processor requires a computational effort more than an order of magnitude less than the MVDR processor is a very important point that has great bearing upon any processor intended for at-sea applications. Space and weight limitations aboard ships often impose severe limitations upon the amount of signal processing that can be conducted.

III. BEAMFORMER SIMULATOR

The simulator designed by the authors allows any number of (spatially coherent) plane wave clutter noise sources to be input at user-specified arrival angles, center frequencies, bandwidths, and amplitudes. Currently, a background noise field with Gaussian statistics is implemented in the simulator,

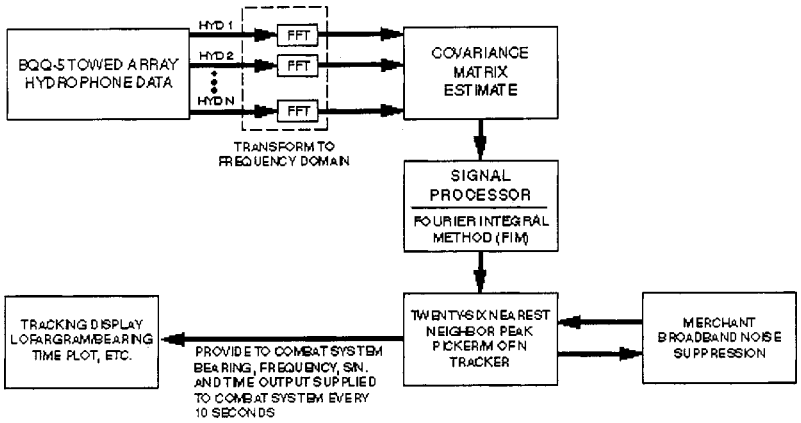


FIG. 1. Inverse beamformer functional block diagram.

with totally independent sequences of random noise values used for each beamforming run. The background noise field in the simulation, in the absence of clutter sources, is spatially uncorrelated; however, a spatially coherent background noise model is now being evaluated for implementation in the simulator. Even Arctic ice noise data (Wilson, 1983), where there is no surface shipping, displays large magnitudes for the off-diagonal elements of the noise covariance matrix, showing that the background noise field is spatially coherent.

The two nonadaptive, beamforming algorithms used for comparison to ABF are CBF and FIM. No attempt is made to claim that either CBF or FIM is ''optimum'' for realistic, spatially correlated noise fields. Obtaining equivalent performance results with FIM and ABF, for example, is a result of high practical importance, because ABF requires at least an order of magnitude more processing power than FIM (e.g., 15 boards versus 1). The simulator results (in Sec. V) will show that FIM outperforms CBF significantly and is comparable to ABF. This means that ABF is not always effective in suppressing clutter for the case of finite, tactical array averaging times and realistic noise conditions at VLF.

It is assumed that the properties of CBF are well known, and will not be reviewed in this article. The properties of FIM (Wilson, 1983; Nuttall and Wilson, 1991; Wilson, 1995; Fabre and Wilson, 1995) have been described elsewhere. FIM is a nonadaptive, covariance matrix (i.e., frequency domain) beamforming algorithm that has a narrower beamwidth than CBF by a factor of $\frac{2}{3}$, while a generalized version of FIM has equivalent peak sidelobe levels. The FIM beampattern advantages over CBF result in a 3-dB array gain (AG) advantage over CBF, which results in much lower minimum detectable levels (MDLs) at the output of the detector (Fabre and Wilson, 1995). FIM is the beamforming algorithm for inverse beamforming (IBF) whose processing schematic is shown in Fig. 1. The IBF postprocessor, namely the 26 nearest-neighbor peak picker (TSNNPP) and M of N tracker, will not be discussed in detail in this article; however, the optimum M of N values, as determined from at-sea data processing, use 24 of 36 averaging time epochs (updated every 10 s). Therefore, the number of time epochs (i.e., the number of independent runs) used in the simulation was taken as 25. The high variability of individual simulation runs shows the realism in using narrow-band spatially correlated interference noise as well as background noise when trying to detect very weak signals. The presentation of 25 overlapping curves under the same conditions gives the reader a chance to observe the variability of a processing technique.

One could misinterpret the significance of the plots of 25 runs in our figures by asserting that averaging the output power over 25 runs increases the number of degrees of freedom in their effective covariance matrices. That is incorrect and disallowed; rather, the estimated CSDMs for each run have not been modified in any way whatsoever. Superposition of the 25 output runs allows the reader to easily investigate the mean, bias, and standard deviation of each processor directly at the beamformer output. If one processor has better performance at this output comparison point, and if all processors have the same identical inputs, that processor is deemed better, at least under those conditions. Whether one processor is linear or nonlinear is totally irrelevant.

It must be emphasized that we are not averaging 25 runs, which would require using 25×20 s$=500$ s of data to generate a single beamformer output; rather, we are using just 20 s for our individual observation time T, and then doing 25 such separate statistically independent runs, for purposes of establishing valid mean and standard deviation comparisons. The restriction to 20-s epochs per run is done on purpose, in order to mimic the practical scenario where contacts move quickly across the field of observation. Both processors must operate on the same 20-s epochs if proper account is to be taken of the nonstationary pressure field in practice.

If one were to assume an ideal but totally unrealistic case, where the acoustic environment was perfectly stationary over all 25 runs (25×20 s/run$=500$ s), then nonlinear processors, such as MVDR and MUSIC, would perform significantly better. It can be shown for linear processors, such as CBF and FIM, that the average of 25 runs of 20 s each is mathematically equivalent to a single run of 500 s. The same property does not hold for nonlinear processors, such as MVDR or MUSIC. Therefore, one should not equate the average of 25 runs for nonlinear processors with the result that would be obtained by averaging for 500 s. It is the high time-nonstationarity of the ocean noise environment that limits the averaging time to 20 s or less. Averaging over 500 s in a nonstationary environment, of course, degrades the performance of all processors significantly.

IV. ABF ALGORITHMS EVALUATED

The objective of this section is to provide a simple explanation of how and where the ABF algorithms are nonlinear, and, thus, noisy when the covariance matrix cannot be estimated exactly in a finite-averaging interval. The basic equations of the MVDR ABF algorithm (Burdic, 1991) are given by Eqs. (1)–(4) below. The $M\times 1$ steering vector is, for temporal frequency f,

$$D_f(u)=\left[1\ \exp\left(i2\pi f\frac{d}{c}u\right)\cdots\exp\left(i2\pi f(M-1)\frac{d}{c}u\right)\right]^T, \tag{1}$$

where $u=\cos\theta$ and θ is the angle measured from endfire of the equispaced line array with M elements at spacing d. The estimated $M\times M$ spectral covariance matrix is

$$\hat{G}_f=\begin{bmatrix}\hat{G}_f(0) & \hat{G}_f(d) & \cdots & \hat{G}_f((M-1)d)\\ \hat{G}_f(-d) & & & \\ \vdots & & & \\ \hat{G}_f((1-M)d) & \cdots & & \hat{G}_f(0)\end{bmatrix}, \tag{2}$$

where $\hat{G}_f(md)$ is the Toeplitz-averaged estimated cross spectrum at frequency f and physical separation md. The resulting spatial directionality estimate at frequency f is

$$\hat{N}_f(u)=\frac{1}{D_f^H(u)\hat{G}_f^{-1}D_f(u)}, \tag{3}$$

where H denotes the complex conjugate transpose. Finally, the optimum weight vector for each u is

$$W_f(u)=\frac{\hat{G}_f^{-1}D_f(u)}{D_f^H(u)\hat{G}_f^{-1}D_f(u)}. \tag{4}$$

Mathematically, the estimated spatial correlations are grouped into an $M\times M$ spectral covariance matrix according to Eq. (2), inverted, then pre- and postmultiplied by the steering vector according to Eq. (3), and the resultant scalar inverted again, yielding a directionality estimate. However, the requirement to estimate an inverse matrix from limited time samples of data can cause performance degradation and thwart any claim to optimality for ABF that accompanied the *known* correlation case.

The MVDR processor is not allowed the use of CSDM estimates of high quality in tactical conditions. Fast moving contacts and low input SNRs frequently limit averaging times to less than 20 s, which leads to rather poor-quality covariance estimates. In fact, this is the reason that MVDR typically resorts to eigendecomposition followed by *ad hoc* modification of the smallest eigenvalues, in an attempt to get a fairly reasonable inverse covariance matrix. Unfortunately, this highly nonlinear procedure leads to small-signal suppression and loss of detection of some weak targets.

A theoretical analysis of the highly nonlinear MVDR processor is not practical for measured data, nor would it help to merely evaluate the partial derivatives of the processor output with respect to the individual terms in the estimated covariance matrix. The individual elements of the covariance matrix are highly statistically dependent on one another; after all, they are all using exactly the SAME limited amount of element data, although in slightly different nonlinear forms. To try to take this dependency into account by a theoretical treatment, which must be done for a proper analysis, is impractical. Also, that analysis would undoubtedly have to be limited to Gaussian input processes, to make it anywhere near tractable. Instead, our simulation approach allows for arbitrary input processes and statistics.

A. The enhanced MVDR (EMVDR) algorithm

Norman Owsley of the Naval Undersea Warfare Center (NUWC) Division, Newport, RI (Owsley, 1998) provided the algorithms for two specific versions of MVDR in current use. Enhanced MVDR (EMVDR) is obtained by using the Toeplitz-averaged (TA) estimated covariance matrix for a smaller subarray in Eq. (2) above, and then performing the following eigenvalue decomposition on the TA covariance matrix, G_f^{TA}:

$$G_f^{\mathrm{TA}}=\sum_{m=1}^{M}\lambda_m U_m U_m^H. \tag{5}$$

Here, $I=UU^H=U^HU$, and λ_m is the mth eigenvalue of G_f^{TA} for an M-element equally spaced line array. U_m is the mth eigenvector corresponding to the eigenvalue λ_m, and $U=[U_l\ldots U_m]$. Assuming the eigenvalues are magnitude ordered from high ($m=l$) to low ($m=M$) values, the largest N_d or ''dominant'' eigenvalues are used in EMVDR as follows:

$$G_g^{\mathrm{TA}}=e\sum_{m=1}^{N_d}(\lambda_m-\lambda_0)U_m U_m^H+\lambda_0 I, \tag{6}$$

where ''average eigenvalue'' λ_0 is

$$\lambda_0=\frac{1}{M-N_d}\sum_{m=N_d+1}^{M}\lambda_m. \tag{7}$$

The enhancement factor e and the dividing line N_d between the dominant and the remaining eigenvalues are chosen empirically from past processing of at-sea data. Typical values were provided by Norman Owsley (Owsley, 1998). The resulting eigenvector decomposition for EMVDR can be expressed as

$$G_g^{\mathrm{TA}}=\sum_{m=1}^{M}\hat{\lambda}_m U_m U_m^H, \tag{8}$$

where modified eigenvalues

$$\hat{\lambda}_m=\begin{cases}e\lambda_m+(1-e)\lambda_0 & \text{for } 1\leqslant m\leqslant N_d\\ \lambda_0 & \text{otherwise.}\end{cases} \tag{9}$$

It is noted that inverse matrix

$$(G_g^{\mathrm{TA}})^{-1}=\sum_{m=1}^{M}\frac{1}{\hat{\lambda}_m}U_m U_m^H \tag{10}$$

constitutes a highly nonlinear transformation on the measured array data.

In the simulation results in the next section, a 48-element ($M=48$), equally spaced line array is used. The

above formulation of EMVDR would have required prohibitively large computational power, especially in performing a 48 by 48 singular value decomposition (SVD) calculation for each frequency bin every 10 or 20 s. To overcome this numerical limitation, the 48-element array was divided into 6 subarrays with 8 elements each, and the EMVDR weights were calculated based on 6 subarrays, each prebeamformed in the broadside direction. The obvious disadvantage of this subarray EMVDR technique is that the preformed broadside beams for the subarrays somewhat degrade the ''optimal weights'' for MVDR given by Eq. (4) in their ability to suppress clutter, or ''notch out'' noise contributions from surface ships. However, a great computational advantage is achieved in performing the SVD calculation only for a 6×6 complex matrix, rather than for a 48×48 complex matrix.

The FFT (or averaging) time used in the simulations is 20 s with 50% FFT overlap and Hann temporal weighting. This results in a 0.05-Hz frequency resolution with a 10-s update interval. To make the EMVDR calculation statistically stable, the weights in Eq. (4) were calculated from 39 1-s FFTs with 50% overlap, thereby fitting within the same 20-s averaging time. One obvious disadvantage of using an ABF algorithm for short tactical sonar update intervals is that the optimal weights must be calculated with a coarse resolution (1 Hz/l s FFTs in this case) relative to the beamforming FFT (0.05 Hz in this case) in order to achieve stability for EMVDR or ABF in general. The above limitation poses the question: Can optimum weights calculated at 1-Hz resolution for ABF efficiently reject noise in the beamformer with 0.05-Hz bandwidth? If longer averaging times are attempted for ABF, the acoustic field becomes non-time-stationary, because surface shipping sources with nonzero bearing rates pass from one beam into an adjacent beam over these longer averaging periods.

B. The MUSIC algorithm

Once EMVDR is calculated, as specified in the previous section, the multiple signal classification (MUSIC) output (Schmidt, 1986) can be calculated with little additional computation:

$$\mathrm{MUSIC}(u)=\left(M-\sum_{m=1}^{N_d} |D^H(u)U_m|^2 \right)^{-1}. \tag{11}$$

MUSIC exploits the noise subspace and realizes a ''small'' value in the denominator in Eq. (11) for directions ($u=\cos\theta$) where there are plane wave arrivals.

C. Vernier MVDR

Although the ''optimum'' complex adaptive weights given by Eq. (4) were computed at 1-Hz frequency resolution, they can be additionally combined for Vernier FFT beamforming resolution (0.05 Hz/20-s averaging time) to suppress clutter. At this stage of the processing, the adaptive weights at 1-Hz resolution do not constitute an optimum filter to minimize noise power out of a beamformer with 0.05-Hz resolution, but they can be used to develop an approximation. Define

$$Y_f(u,j)=W_f^H(u)V_f(j) \quad \text{for } 1\leqslant j\leqslant N_p, \tag{12}$$

where $V_f(j)$ are the complex conventional beam time series for epoch j_i and N_p is the number of pieces (time segments) in the 20-s averaging time. Forming the DFT and Hann weighting the time series at epoch j with sequence $\{B_j\}$ yields

$$Z_f(u)=\sum_{j=1}^{N_p} \exp(-i2\pi f N_{fj}\Delta t/2)Y_f(u,j)B_j, \tag{13}$$

where Δt is the sample interval and N_f is the DFT size. One then computes power function $|Z_f(u)|^2$ to form the Vernier MVDR FRAZ output as a function of angle u and frequency f.

D. CBF and FIM

CBF and FIM are computed as follows; the noise directionality estimate is

$$\hat{N}_f(u)=\frac{fd}{c}\sum_{m=1-M}^{M-1} \exp(-i2\pi fmdu/c)\hat{G}_f(md)w_m, \tag{14}$$

where d is the array element separation, M is the number of elements, c is the speed of sound, and weightings

$$w_m=\begin{cases}(M-|m|)/M & \text{for CBF}\\ 1 & \text{for FIM}\end{cases}$$
$$\text{for } 1-M\leqslant m\leqslant M-1. \tag{15}$$

FIM's beamwidth is approximately $\frac{2}{3}$ narrower than CBF's beamwidth, and FIM has 3 dB more array gain than CBF (Nuttall and Wilson, 1991; Wilson, 1995). Note that FIM and CBF are equivalent computationally, and that both FIM and CBF require far less (by orders of magnitude) computational power than any of the ABF algorithms. Previous publications (Wilson, 1983; Nuttall and Wilson, 1991; Wilson, 1995; Fabre and Wilson, 1995; Wilson and Veenhuis, 1997) contain much more detail on the IBF algorithms; the only new FIM technique presented in this article is the generalized FIM (GFIM) weighting defined as

$$w_m=1-\cos(2\pi m/M) \quad \text{for } 1-M\leqslant m\leqslant M-1. \tag{16}$$

A bad feature of the FIM beampattern relative to the CBF beampattern is FIM's higher peak sidelobe levels at frequencies above VLF. GFIM eliminates this problem by weighting the widest element separations in the covariance-matrix low relative to the intermediate hydrophone pair separations. In fact, GFIM zero-weights the hydrophone self-pairs (zero separation), and the weighting increases with hydrophone pair separation until a separation of one-half the total array aperture is reached. The GFIM weighting then decreases until maximum separation is reached. Figure 2 shows a comparison of the CBF, FIM, and GFIM beampatterns at VLF for a 48-element, equally spaced line array. GFIM is the beamformer that performed best on the measured data performance comparison between FIM and MVDR (Solomon *et al.*, 1999) in an independent test by Australian scientists. GFIM was developed by one of the authors and is the same

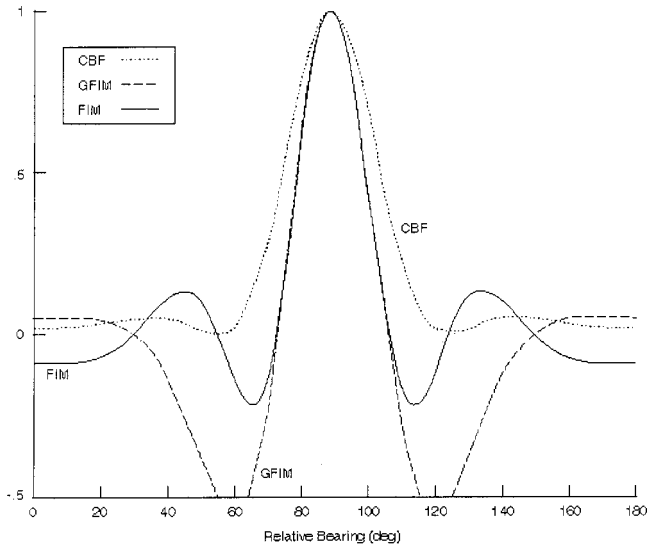


FIG. 2. Estimate of acoustic field directionality for broadside arrival at a 48-element, equally spaced, line array at 4% design frequency.

beamforming algorithm designated WFIM in Solomon *et al.* (1999).

One may correctly note that the covariance matrix estimates at the longer lags (separation distances) have less stability than those for the shorter lags. However, one should not fall into the usual trap of assuming that the longer lag estimates should necessarily be deemphasized. Instead, one must always consider the inherent trade-off that takes place between angular resolution and stability. The (less stable) longer lags contain most of the array's angular resolution capability that MUST be exploited at VLF. It is frequently better to have noisy indications that there are TWO weak close targets than to get a single stabler (misleading) indication that only one arrival is present. Furthermore, additional posteamformer processing can track and separate the two separate indicators, which is not a possibility for a single indicator. Also, the covariance estimates used by FIM are unbiased, whereas those employed by CBF and GFIM are biased.

E. A simple example of the impact of ABF's nonlinearity

This article uses realistic simulations in the next section to show that FIM outperforms CBF significantly, and that FIM's performance is slightly better than the best of the ABF algorithms used in this simulation. This result is achieved even though ABF requires orders of magnitude more processing power. In practical terms of a VME, commercial off-the-shelf (COTS) computer, ABF requires 10 to 15 times more processing VME boards than FIM to achieve a performance level comparable to FIM. This is a dramatic demonstration of a drawback of ABF for short averaging times, due to the complexity of the eigenvalue/eigenvector ABF algorithms, which magnify errors made in the estimation of the covariance matrix. In this section, the nonlinear impact of dividing by the inverse of the covariance matrix is demonstrated by a very simple example. For a two-element, equally spaced line array of spacing d, the 2×2 covariance matrix is given by

$$\hat{G}_f=\begin{bmatrix}\hat{G}_f(0) & \hat{G}_f(d)\\ \hat{G}_f^*(d) & \hat{G}_f(0)\end{bmatrix}, \tag{17}$$

where the off-diagonal elements are complex. Some straightforward algebra yields a MVDR directionality estimate at broadside

$$[D_f^H(u)\hat{G}_f^{-1}D_f(u)]_{u=0}^{-1}=\tfrac{1}{2}[\hat{G}_f(0)+\hat{G}_{fr}(d)]-\frac{1}{2}\left[\frac{\hat{G}_{fi}^2(d)}{\hat{G}_f(0)-\hat{G}_{fr}(d)}\right], \tag{18}$$

where broadside steering vectors are assumed [Eq. (1)] and where the subscripts r and i refer to the real and imaginary parts of the complex number, respectively.

Small errors in the linear difference term, $\hat{G}_f(0)-\hat{G}_{fr}(d)$, in the denominator of the nonlinear term in Eq. (18) can have a large impact on the ABF output. By contrast, FIM and CBF are both linear in all the elements $\{\hat{G}_f(md)\}$ of the covariance matrix.

For a three-element, equally spaced line array, Eq. (3) yields (for the broadside beam, $u=0$)

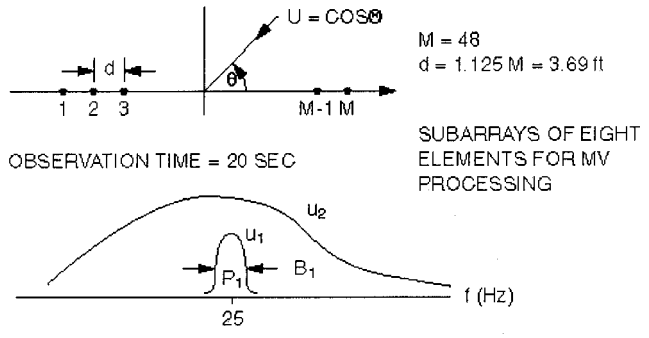
$$\left[\frac{1}{D_f^H(u)\hat{G}_f^{-1}D_f(u)}\right]_{u=0}=\frac{\hat{G}_f^2(0)+\hat{G}_f(0)\hat{G}_f(2d)-2\hat{G}_f(d)^2}{3\hat{G}_f(0)-4\hat{G}_f(d)+\hat{G}_f(2d)}. \tag{19}$$

Again, small errors in the denominator of Eq. (19) can lead to large fluctuations in the ABF output. For example, if $\hat{G}_f(2d)=4\hat{G}_f(d)-3\hat{G}_f(0)$, the output becomes infinite.

Thus, ABF's objective of ''nulling'' two or more line clutter sources (not in the beam direction of interest) requires a very accurate estimate of the covariance matrix, which can only be obtained through use of large averaging times. For the real-ocean environment containing highly fluctuating noise fields caused by interfering multipath arrivals from dynamic shipping sources, ABF's performance for short-averaging times will be degraded, as shown by the results in the next section. Operations at VLF provide fundamental limitations on the performance of ABF (as discussed in Sec. II), in addition to the limitations caused by estimation errors for the covariance matrix.

V. SIMULATION RESULTS

Numerous simulations with a wide variety of noise parameters have been run; the limited number of results shown in this article are typical of all the runs performed. The only reasonable practical method to compare FIM and MVDR, in a fashion that correctly and exactly accounts for all the nonlinear transformations of the MVDR processor, is to conduct a direct simulation comparison of the two processors under investigation, using the ''same playing field''. By this is meant that the same TOTAL observation time T and the same TOTAL array extent L must be made available to both processors. It does not in any way mean that both processors must use the same processing bandwidths. This is a very



- Arrival 1: $u_1 = -0.4$, $P_1 = -37$ dB, $B_1 = 0.07$ Hz, $f_1 = 25$ Hz
- Arrival 2: $u_2 = +0.4$, $P_2 = -30$ dB, $B_2 = 0.56$ Hz, $f_2 = 25$ Hz
- FIM: One FFT over one 20-sec segment: 0.05 Hz resolution

MVDR:: 15 2.5-second segments with 50% overlap: 0.4-Hz resolution. N_d = two dominant eigenvalues, and an enhancement factor e = 10.

FIG. 3. Simulation parameters.

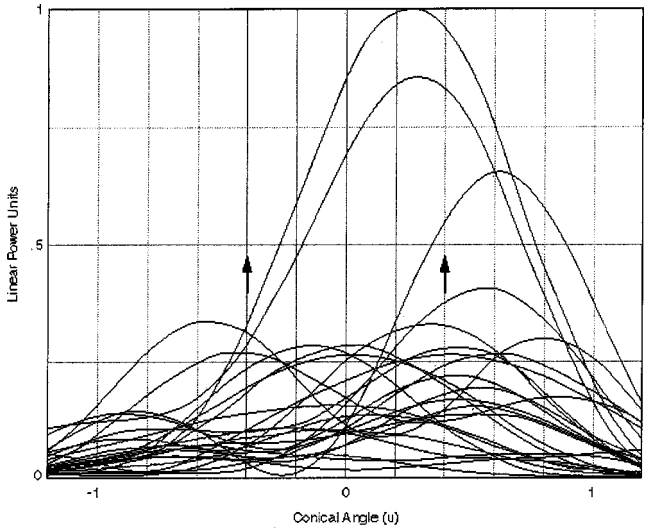


FIG. 4. Twenty-five runs of CBF.

important point: how two different processors choose to utilize their fundamentally limited resources, namely T and L, is entirely up to the systems designers. Each processor can optimize its own choice of bandwidths, time constants, etc., in total disregard for other competing processors. The only valid comparison point is at the system output; for example, how well did they both do in detecting a -35-dB weak signal in the presence of a -20-dB nearby interference, which overlapped in frequency? A recent performance comparison of FIM and MVDR (Solomon *et al.*, 1999) using measured data supports this conclusion.

Identical SNRs were employed in the simulation processors. Both system inputs were fed with the same identical set of data, to guarantee a fair comparison, which is consistent with our philosophy of using the ''same playing field.'' The fact that the SNR at some internal point(s) of the MVDR processor may be lower than that for the corresponding point(s) of the FIM processor is an inescapable fact for the MVDR implementation. It is unreasonable to impose artificial constraints upon the two processors; the only restrictions that should be fairly allowed are identical conditions at their inputs. A demand that ''all processors should use the same bandwidth and covariance matrix estimates'' is not logical. Each processor should be able to do entirely what it chooses. The only constraints are T, L, input SNRs, and identical data, namely common input information. If one processor has to evaluate a large number of parameters in order to perform adequately, that should not constitute a limitation on a competing procedure which can get by with fewer estimation requirements.

The particular simulation parameters used in this article are indicated in Fig. 3. The line array has 48 elements with separation of 1.125 m and epoch duration of 20 s. Two plane wave arrivals with SNRs of -37 and -30 dB relative to the single-phone noise level are present at each receiving element. The center frequencies of both narrow-band arrivals, as well as the analysis frequency of interest, are 25 Hz.

The FIM employs a single FFT over the 20-s epoch duration, resulting in 0.05-Hz resolution. The MVDR procedures obtain some stability, at the expense of a broader resolution of 0.4 Hz, by employing 15 2.5-s FFTs with 50% overlap. The three MVDR procedures utilize six subarrays of eight elements each and use two dominant eigenvalues ($N_d = 2$) along with an enhancement factor of 10 ($e = 10$). Exactly the same data were used for the simulation comparisons of FIM, CBF, and MVDR.

The estimated directionality is plotted in linear power units versus angle parameter u. The actual arrival angles of the two weak plane waves are indicated by the pair of arrows. Figure 4 gives the results of 25 runs of CBF, which utilizes triangular covariance weighting $\{w_m\}$. Most of the runs yield a single peak in angle; however, notice the impact of a realistic noise field by the large variation in the angular location of the peaks in each of the 25 runs. This lack of resolution for the CBF is confirmed by the sample average of the 25 runs shown in Fig. 5. The arrival angle of the stronger plane wave, simulating a surface ship, is correctly indicated, whereas there is no apparent indication of the weaker plane wave arrival.

The 25 runs computed for the FIM beamforming proce-

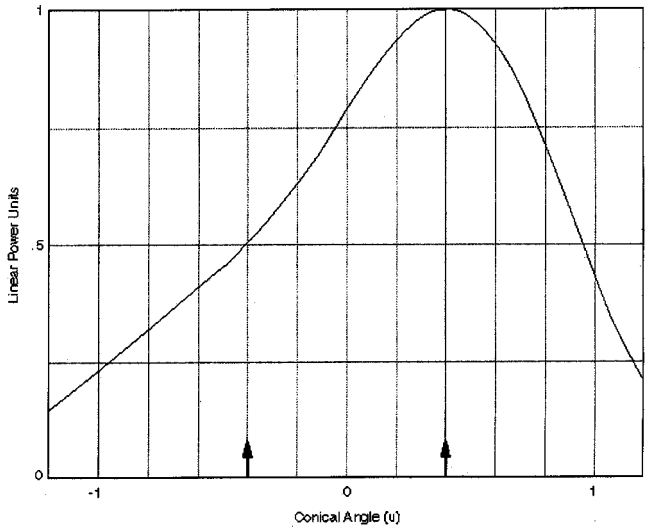


FIG. 5. Sample average of 25 runs for CBF.

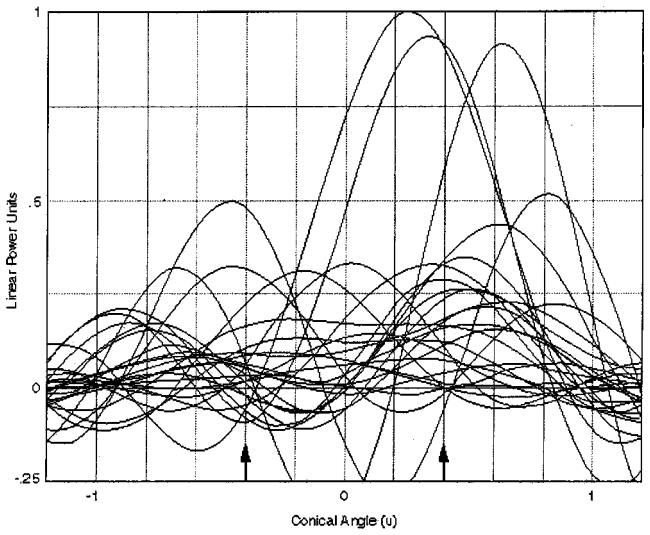


FIG. 6. Twenty-five runs of FIM.

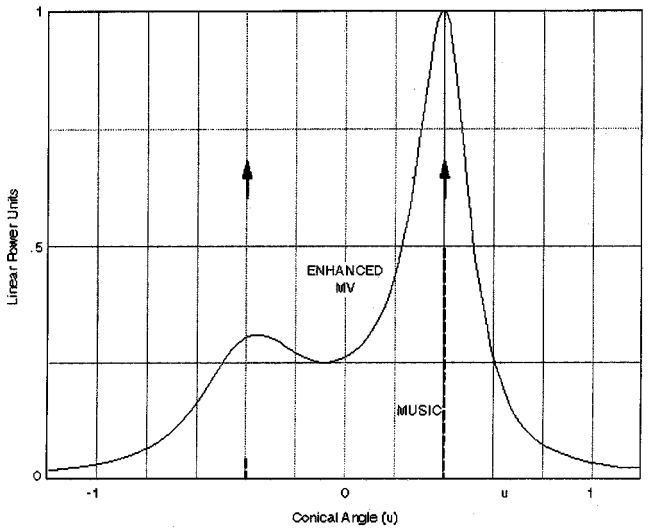


FIG. 8. The MVDR directionality estimates for known correlations—simulating long averaging times.

dure are shown in Fig. 6. Numerous runs experience two peaks within the plotted angular region, due to the narrower FIM beamwidth. The occasional negative directionality estimates are a manifestation of using flat covariance weights $\{w_m\}$, but are of no concern. The sample average of the 25 FIM runs shown in Fig. 7 displays a separate peak in the neighborhood of the weaker arrival, in contrast to the CBF results above.

If the spatial correlation were known exactly and did not have to be estimated, the directionality estimates for EMVDR and MUSIC would be as indicated in Fig. 8. Both procedures correctly indicate the presence of the weaker arrival. However, when the spatial correlations must be estimated from the same 20-s data segment, sample runs for the EMVDR processor display considerable variability, as shown in Fig. 9. The corresponding 25 runs for MUSIC shown in Fig. 10 give only a few indications of the weaker arrival. There is also considerable variability in the estimate of the stronger arrival. The 25 results for the Vernier MVDR procedure shown in Fig. 11 have a great deal of variability, although a number of runs do correctly locate the weaker arrival. The linear average of the 25 runs in Figs. 9–11 are shown by the cross-hatched curve.

Figure 12 shows the sample averages of the 25 runs for each of the three MVDR processors. Based on these averages, the Vernier MVDR technique appears to offer the best performance of the three MVDR processors considered, at least for the parameter values considered. On the average, the indications of the weaker arrival for Vernier MVDR are comparable to, or slightly worse than, those obtained by means of the FIM approach. Additional simulation comparisons are currently underway for other parameter values, for larger total observation times of the order of 2 min and longer. In the case of longer averaging times, the two plane wave arrivals will be stepped in bearing to simulate realistic bearing rates for tactical sonar systems, and the relative performance of nonlinear and linear processors will be addressed.

The importance of the single epoch with short averaging times of 10 to 20 s is that the estimated covariance is not

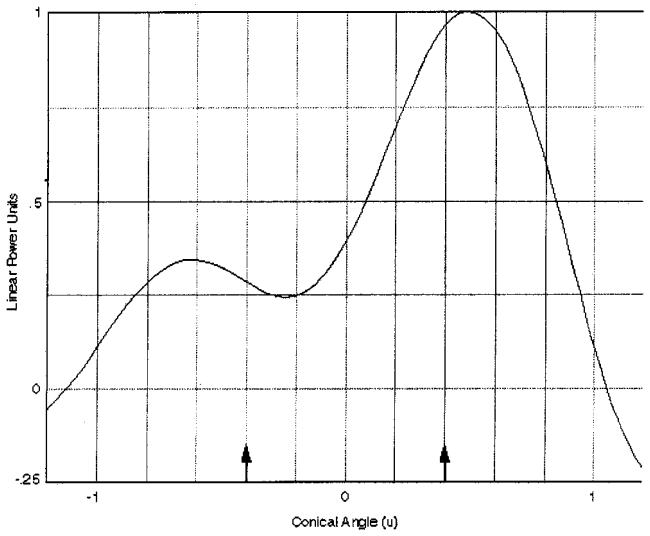


FIG. 7. Sample average of 25 runs for FIM.

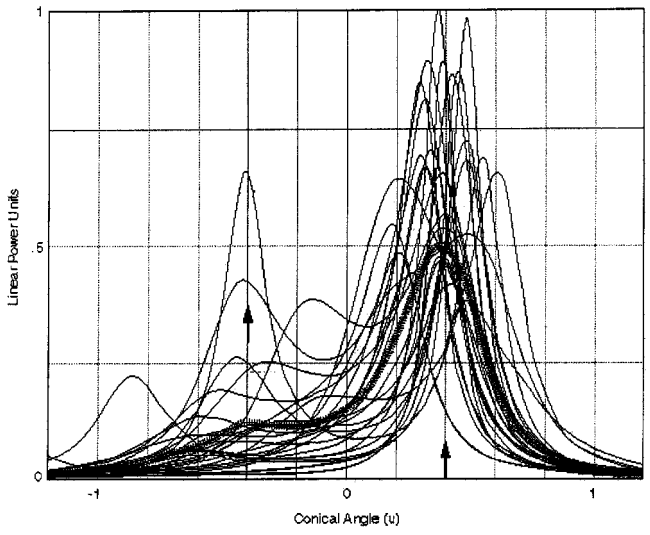


FIG. 9. Twenty-five runs of enhanced MVDR.

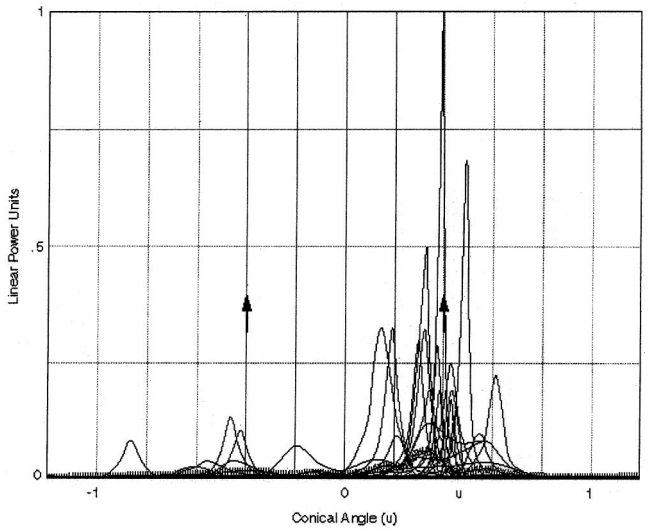


FIG. 10. Twenty-five runs of MUSIC.

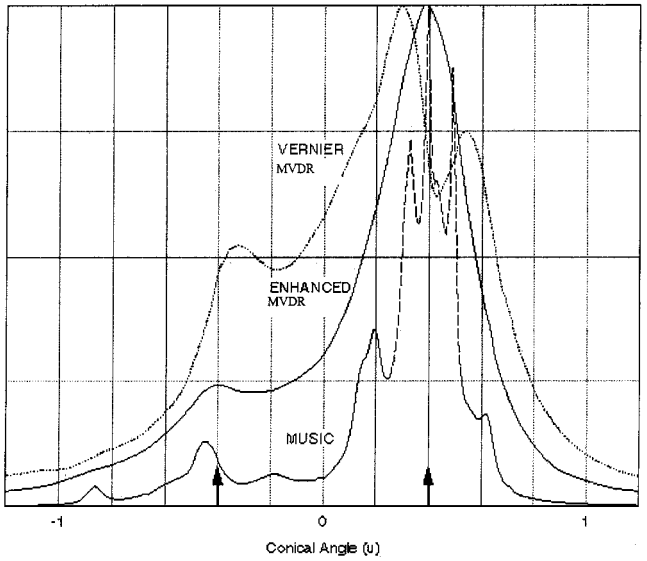


FIG. 12. Sample average of 25 runs for three MVDR processors.

equal to the exact covariance matrix. The results of our simulations show that the error in estimation (the exact covariance minus the measured covariance matrix) degrades performance much more significantly for nonlinear algorithms like MVDR than for linear algorithms like FIM. The average of 25 single epoch covariance matrices would not be used in practical sonars, and a peak picker and tracker would be used instead. The average covariance matrix contains both the signal and noise, and the results show that in averaging, the noise is suppressed and the signal-to-noise ratio is enhanced. Naturally, one cannot average for 8 min (25 epochs) in tactical situations, but the average is indicative of how a tracker might perform over 8 min (25 epochs).

For comparisons of beamformers, the bias and variability of their outputs under identical conditions is an excellent basis on which to make recommendations. We stress that histograms of peak bearings would be a worthwhile addition to our results in future simulations. It is not realistic to generate the ROCs pertaining to the beamformer outputs, because in IBF, the beamformers are followed by peak pickers, trackers, etc., which differ significantly from one system realization to another. The results we have presented are ample to point out an important fundamental limitation of MVDR; furthermore, these limitations are manifest in the FIM and MVDR side-by-side performance comparisons using measured data (Solomon *et al.*, 1999).

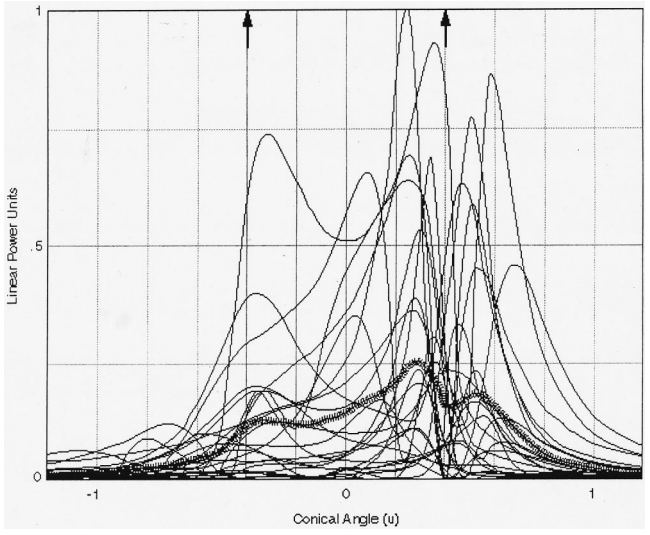


FIG. 11. Twenty-five runs of enhanced Vernier MVDR.

VI. CONCLUSIONS

The simulation results show that the super-resolution capabilities of the ideal element-based MVDR approach are not realized on the average when the available data is of short duration and the input SNR is low, although several runs do show two peaks. This is due to the fact that the estimates of the pertinent covariance matrices and their inverses are highly nonlinear and inherently noisy, thereby adversely affecting performance. In addition, the MVDR approach has fundamental limitations at VLF. The FIM approach, although it does not possess any super-resolution capability, is a robust technique capable of detecting low-level arrivals in realistic noise environments, and has no fundamental limitations at VLF. Furthermore, FIM has better resolution capability than CBF and outperforms CBF by a significant margin. The difficulties with MVDR processors using short averaging times at low SNR are not well known. Also, no direct comparison of the signal detection capability had ever been made between the MVDR and the FIM beamformers prior to the Australian Navy (Solomon *et al.*, 1999) and our results. The fact that the two processors perform nearly equally at low SNR and for limited observation times is an important one for the engineering community, so that unwarranted additional effort is not devoted to a technique with limited capability, namely MVDR. A very recent publication (Solomon *et al.*, 1999) makes a side-by-side performance comparison of FIM, MVDR, and CBF, using measured data, and FIM outperformed both MVDR and CBF by a wide margin. This confirms that our simulations are on the conservative side, and that FIM can outperform MVDR when real sea-data is encountered.

Burdic, W. S. (**1991**). *Underwater Acoustic System Analysis*, 2nd ed. (Prentice–Hall, Englewood Cliffs, NJ), pp. 192–203.
Capon, J. (**1969**). ''High-Resolution Frequency-Wavenumber Spectrum Analysis,'' Proc. IEEE **57**, 1408–1418.
Donald, J. B., Nuttall, A. H., and Wilson, J. H. (**2000**). ''Adaptive Beamforming for Noise and Interference Suppression,'' *Undersea Defence Technology Symposium*, Sydney, 7–9 February 2000, pp. 192–200.
Fabre, J. P., and Wilson, J. H. (**1995**). ''Minimum detectable level evaluation of inverse beamforming using Outpost Sunrise data,'' J. Acoust. Soc. Am. **98**, 3262–3278.
Kneipfer, R. (**1997**). ''Algorithm Description for a Beam Space, Minimum Variance Soft Constraint (MVSC) Adaptive Beamformer,'' NUWC-NPT Technical Report 10, 682.
Mohnkern, G. L. (**1989**). ''Effects of Errors and Limitations on Interference Suppression,'' *NOSC Technical Document 1478*, Naval Ocean Systems Center.
Nuttall, A. H., and Hyde, D. W. (**1969**). ''Unified Approach to Optimum and Suboptimum Processing for Arrays,'' *USL Report Number 992*, Naval Underwater Systems Center, New London, CT, 22 April 1969.
Nuttall, A. H., and Wilson, J. H. (**1991**). ''Estimation of the acoustic field directionality by use of planar and volumetric arrays via the Fourier integral method,'' J. Acoust. Soc. Am. **90**, 2004–2019.
Owsley, N. (**1998**), personal communications.
Schmidt, R. O. (**1986**). ''Multiple Emitter Location and Signal Parameter Estimation,'' IEEE Trans. Antennas Propag. **AP-34**(3), 276–280.
Schuster, J. G. (**1998**). ''Review of IBF,'' Chief of Naval Operations Letter N875/8U657523 of 14 December 1998.
Solomon, J. S. D., Knight, A. J., and Greening, M. V. (**1999**). ''Sonar Array Signal Processing for Sparse Linear Arrays,'' Defense Science and Technology Organization (DSTO) Technical Report June 1999.
Thompson, A. R., Moran, J. M., and Swenson, G. W. (**1994**). *Interferometry and Synthesis in Radio Astronomy* (Krieger, New York).
Vural, A. (**1974**). ''Adaptive Beamforming,'' course presentation.
Wilson, J. H. (**1983**). ''Signal detection and localization using the Fourier series method (FSM) and cross-sensor data,'' J. Acoust. Soc. Am. **73**, 1648–1656.
Wilson, J. H. (**1995**). ''Applications of inverse beamforming (IBF),'' J. Acoust. Soc. Am. **98**, 3250–3261.
Wilson, J., and Veenhuis, R. **1997**). ''Shallow water beam forming with small aperture, horizontal towed arrays,'' J. Acoust. Soc. Am. **101**, 384–394.

The mode-coupling Liouville–Green approximation for a two-dimensional cochlear model

Lloyd Watts[a)]
500 Chiquita Avenue, #2, Mountain View, California 94041

(Received 28 December 1999; accepted for publication 10 July 2000)

The Liouville–Green [or Wentzel–Kramers–Brillouin (WKB)] approximation for the two-dimensional cochlear mechanics problem disagrees with the finite-difference solution in the region after the response peak. This disagreement has left doubts about the validity of the Liouville–Green approximation, and has never been satisfactorily explained. In this paper, it is shown that the Liouville–Green approximation fails to satisfy Laplace's equation. A new solution is proposed, called the *mode-coupling Liouville–Green approximation*, in which energy is coupled into a second wave mode, so as to obey Laplace's equation. The new approximation gives excellent quantitative agreement with the finite-difference solution. Furthermore, it may provide an explanation for a second vibration mode observed in biological cochleas. Also proposed is a high-order formulation of the stapes displacement term, which is necessary to obtain good agreement between the Liouville–Green approximation and finite-difference solutions at low frequencies. © *2000 Acoustical Society of America.* [S0001-4966(00)05010-4]

PACS numbers: 43.64.Bt, 43.64.Kc [LHC]

I. INTRODUCTION

The Liouville–Green (LG), or WKB, approximation to the cochlear mechanics problem has been studied extensively prior to the mid-1980's by Zweig *et al.* (1976), Steele and Taber (1979), and de Boer and Viergever (1982). However, the advent of fast computers and efficient numerical algorithms has prompted many researchers to abandon the analytic, approximate LG method in favor of ''exact'' numerical methods such as the finite-difference (FD) and finite-element (FE) methods. These exact solutions allow efficient computation of the response to a given stimulus, and sometimes expose unexpected behavior. Unfortunately, they do not provide insights into the complex wave dynamics that can be provided by an accurate analytic expression for the solution.

In 1979, Steele and Taber compared the two-dimensional Liouville–Green approximation to the finite-difference solution of Neely (1981), as shown in Fig. 2. The LG approximation is an excellent approximation to the FD solution up to a few millimeters past the response peak, at which point the two solutions suddenly start to diverge. This discrepancy was addressed directly by de Boer and Viergever (1982); they stressed that the eikonal equation in the LG analysis has multiple roots, and noted that the change in slope in the amplitude response curves corresponds to the sudden emergence of a second wave mode. However, they were unable to predict accurately when the second wave mode would emerge. This uncertainty about the validity of the Liouville–Green approximation also contributed to its falling out of favor in the hearing research community.

In this paper, it will be shown that the Liouville–Green approximation fails to satisfy Laplace's equation. A new approximation, called the mode-coupling Liouville–Green approximation, is proposed which is a linear superposition of the usual traveling wave mode and a second cutoff mode. The amplitude of the second wave mode is determined by requiring that Laplace's equation be satisfied on average in a given vertical slice of the fluid. This procedure leads to an accurate prediction of the emergence of the second wave mode, and produces an approximation that agrees quantitatively with the finite-difference solution. Moreover, the new approximation provides an explanation for a second wave mode first observed in squirrel monkey basilar membrane responses (Rhode, 1971).

In this paper, we also introduce a high-order formulation for the stapes displacement, which is necessary to obtain good agreement between the Liouville–Green approximation and numerical solutions at low frequencies.

II. THE LIOUVILLE–GREEN APPROXIMATION

The two-dimensional cochlear model is shown in Fig. 1, where $\phi(x,y,t)$ is the velocity potential, ρ is the fluid density, h is the height of the duct, L is the length of the duct, and $S(x)$, $\beta(x)$, and $M(x)$ are the stiffness, damping, and mass, respectively, of the basilar membrane (BM). The BM is located at $y=h$. The fluid is assumed to be incompressible.

The Liouville–Green approximation of the 2D cochlear model has been derived previously in the literature (Steele and Taber, 1979), and is the starting point for the present discussion. For brevity, only the major results are shown below.

The approximation for the velocity potential for a sinusoidal input with frequency ω is given by

$$\phi(x,y,t)=\frac{C\omega\cosh(ky)}{\cosh(kh)\sqrt{\tanh kh+kh\,\mathrm{sech}^2\,kh}}\times\exp\left(i\omega t-i\int_0^x k(u)\,du\right), \tag{1}$$

a)Electronic mail: lwatts@lloydwatts.com, www.lloydwatts.com

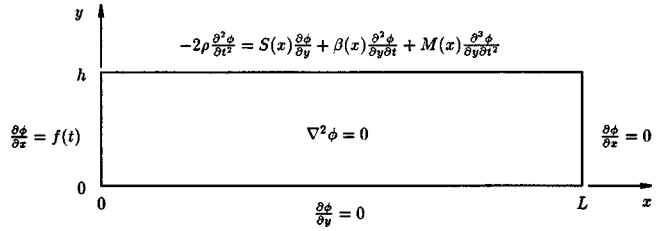


FIG. 1. The two-dimensional cochlear model.

where C is a constant proportional to the stapes displacement, and $k(x)$ is the wave number, which is found by solving the dispersion relation (or eikonal equation)

$$k(x)\tanh k(x)h = \frac{2\rho\omega^2}{S(x)+i\beta(x)\omega - M\omega^2}. \tag{2}$$

The vertical displacement, δ, of the basilar membrane is related to the vertical velocity, $\mathbf{v}_y$, of the fluid at $y=h$, and the velocity potential, by the following relations:

$$\frac{\partial\delta}{\partial t} = \mathbf{v}_y = -\frac{\partial\phi}{\partial y}. \tag{3}$$

A high-order formulation for the stapes displacement d_{st} has been determined by Watts (1992) to be

$$d_{\mathrm{st}} = \frac{C\mathrm{T}(k_0,k_0')\exp(i\omega t)}{h\sqrt{\tanh(k_0h)+k_0h\,\mathrm{sech}^2(k_0h)}}, \tag{4}$$

where

$$T(k_0,k_0') = \tanh(k_0h)\left[1 - \frac{ik_0'}{k_0^2} - \frac{2ik_0'h(1-k_0h\tanh(k_0h))}{k_0(2k_0h+\sinh(2k_0h))} - \frac{ik_0'h\tanh(k_0h)}{k_0}\right] + \frac{ik_0'h}{k_0}, \tag{5}$$

where k_0 and k_0' are the values of k and k', respectively, at $x=0$, and the apostrophe indicates differentiation with respect to x. This high-order formulation for the stapes displacement is the key to obtaining a good agreement with the finite-difference solutions at low frequencies.

Finally, the displacement ratio $D(x,\omega)$ can be shown to be

$$D(x,\omega) = \frac{\delta}{d_{\mathrm{st}}} = ikh\frac{\tanh(kh)}{T(k_0,k_0')} \times\sqrt{\frac{\tanh(k_0h)+k_0h\,\mathrm{sech}^2(k_0h)}{\tanh(kh)+kh\,\mathrm{sech}^2(kh)}} \times\exp\left(-i\int_0^x k(u)\,du\right). \tag{6}$$

The magnitude and phase of the displacement ratio in Eq. (6) are plotted in Fig. 2 for the standard Liouville–Green approximation and the finite-difference solution. For comparison with previously published results, we use the parameters of Neely (1981) as shown in Table I.

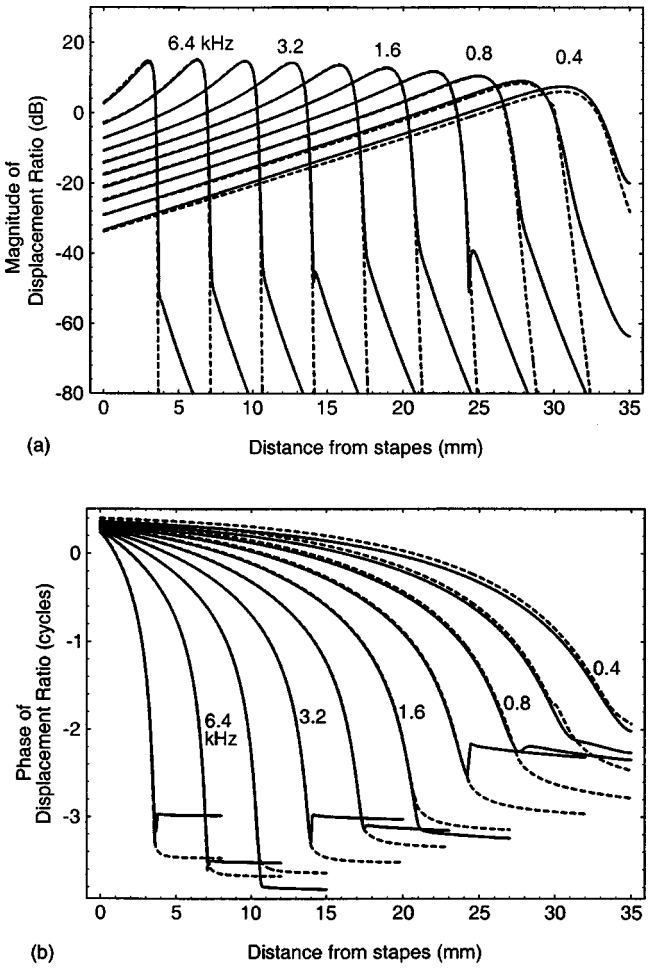


FIG. 2. Comparison of Liouville–Green (dashed) and finite- difference (solid) methods, recomputed after Steele and Taber (1979), and Neely (1981), using the high-order stapes correction of Watts (1992). The two solutions have good agreement up to and just after the peak response. The complete list of frequencies used, in kHz, is (from left to right in the plots): 9.05, 6.4, 4.53, 3.2, 2.26, 1.6, 1.13, 0.8, 0.57, and 0.4.

Note the excellent agreement up to the peak response, and the disagreement just after the peak response.

III. THE LIOUVILLE–GREEN APPROXIMATION FAILS TO SOLVE LAPLACE'S EQUATION

In this section, we shall see that the Liouville–Green approximation associated with the primary root $k(x)$ fails to satisfy Laplace's equation in the fluid just basalward of the cutoff region.

The LG approximation for the velocity potential was designed to satisfy Laplace's equation—that is, to ensure that the flow into any region of space in the x direction is exactly canceled by the flow out of the region in the y direction.

TABLE I. Neely's parameters.

$S(x) = S_0\exp(-x/d)$
$S_0 = 1.0\times10^7$ g s^{-2} mm^{-2}
$\beta = 2.0$ g s^{-1} mm^{-2}
$M = 1.5\times10^{-3}$ g mm^{-2}
$d = 5.0$ mm
$h = 1.0$ mm
$L = 35.0$ mm
$\rho = 1.0\times10^{-3}$ g mm^{-3}

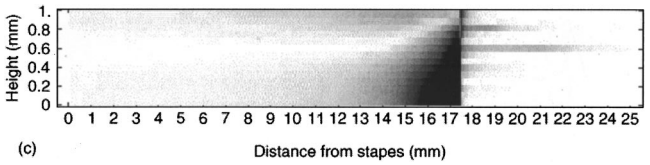


FIG. 3. Relative Laplace error (RLE) shown as a density plot as a function of position in the duct, for Neely's parameters with $f=2.26$ kHz. White corresponds to RLE$=0$; black corresponds to RLE>1.

Since the LG approximation is only approximate, we do not expect that $\nabla^2\phi$ is exactly zero; we expect only that the net accumulation or loss in the region is small compared to the amount flowing through it, for a small relative error. Expressed quantitatively, the solution obeys Laplace's equation if

$$|\nabla^2\phi| \ll |k^2\phi| \quad \text{or} \quad \left|\frac{\nabla^2\phi}{k^2\phi}\right| \ll 1. \tag{7}$$

We shall refer to the term $|\nabla^2\phi/k^2\phi|$ as the *relative Laplace error*, or RLE, and we shall refer to inequality (7) as the RLE *criterion*. A nonzero value of RLE implies that the assumption of fluid incompressibility is being violated.

It can be shown that the RLE criterion reduces to

$$\left|\frac{k'}{k^2}\left(1-\frac{4kh(1-kh\tanh kh)}{2kh+\sinh 2kh}+2k(y\tanh ky-h\tanh kh)\right)\right| \ll 1. \tag{8}$$

Note that the RLE criterion has different properties than the conventional validity criterion $|k'/k^2| \ll 1$. The relationship between the RLE criterion and the conventional validity criterion is discussed in the Appendix.

As an illustrative example, let us arbitrarily use Neely's parameters with an input frequency of $f=2.26$ kHz. In Fig. 3(a), the relative Laplace error is shown as a density plot as a function of position in the duct. The large black region indicates where Laplace's equation is being violated. The violation becomes significant initially at the bottom of the duct at about $x=15$ mm, and rises toward the basilar membrane until about $x=17.4$ mm, which is approximately the location at which the corresponding membrane displacement response in Fig. 2 bends. Thus, it can be seen that the unexpected bend in the membrane displacement response is the result of a gradual process that grows from the bottom of the duct over a 2-mm distance in the x direction. This region corresponds to the void left when the wave makes the transition from long-wave to short-wave behavior, thus ''lifting off'' the bottom of the duct.

Recall that the wave number $k(x)$ is the solution of the dispersion relation, which in fact has infinitely many solutions. Typical wave number trajectories are shown in Fig. 4 for three different frequencies. Each curve illustrates one particular wave mode (corresponding to one particular solution of the eikonal equation). We will define the traveling wave mode as the one whose wave number begins near the origin, and we will define the cutoff mode as the one whose wave number ends near $-i\pi/2$. Note that for the examples of $f=800$ and 1131 Hz, the traveling wave mode wave number encircles the cutoff mode wave number, whereas for $f=400$ Hz, the traveling wave mode is also the cutoff mode.

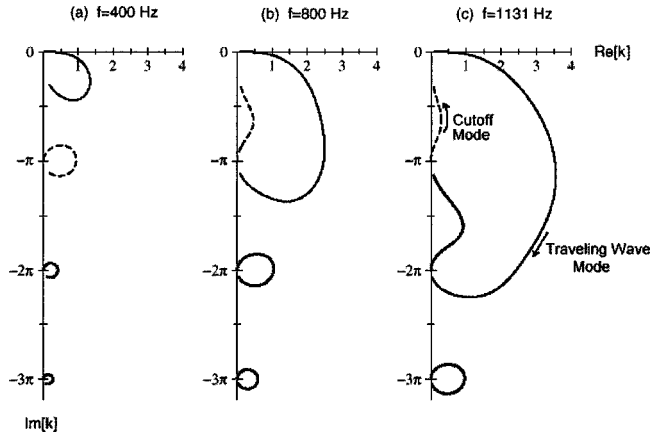


FIG. 4. Example wave number trajectories at three different frequencies, using Neely's parameters. The traveling wave mode begins near the origin. The cutoff mode ends near $-i\pi/2$.

Viergever (1981) observed that the moderate slope in the amplitude and phase responses after the response peak was consistent with the lightly damped cutoff mode. He also observed that the bend in the amplitude curve appeared near the resonance point, so he proposed splicing together the traveling wave and cutoff solutions at the resonance point, prescribing continuity of basilar membrane velocity at the splice. Unfortunately, this procedure resulted in a constant error with respect to the finite-difference solutions.

IV. THE MODE-COUPLING LIOUVILLE–GREEN (MCLG) APPROXIMATION

We now propose the following form of the velocity-potential solution:

$$\phi(x,y,t)=\phi_1(x,y,t)+c(x)\phi_2(x,y,t), \tag{9}$$

where ϕ_1 is the traveling wave solution with wave number k_1, which originates near $k_1\approx 0$ for $x=0$, and ϕ_2 is the traveling wave solution with wave number k_2, which originates near $k_2\approx -(i\pi/2)$ for $x=0$, and $c(x)$ is the coupling coefficient. ϕ_1 and ϕ_2 have been determined already, so all that remains is to determine $c(x)$ such that the composite solution satisfies Laplace's equation.

For Laplace's equation to hold, we must have

$$\nabla^2\phi=\nabla^2\phi_1+c\nabla^2\phi_2+2\frac{\partial c}{\partial x}\frac{\partial\phi_2}{\partial x}+\frac{\partial^2 c}{\partial x^2}\phi_2=0. \tag{10}$$

This equation implies that c must also depend on y to make $\nabla^2\phi(x,y)$ vanish at every point. However, a good approximate solution is possible with $c=c(x)$ alone, so let us specify that the total error must vanish in a vertical slice

$$\int_0^h \nabla^2\phi\,dy=\int_0^h\nabla^2\phi_1\,dy+c(x)\int_0^h\nabla^2\phi_2\,dy+2c'(x)\int_0^h\frac{\partial\phi_2}{\partial x}dy+c''(x)\int_0^h\phi_2\,dy=0. \tag{11}$$

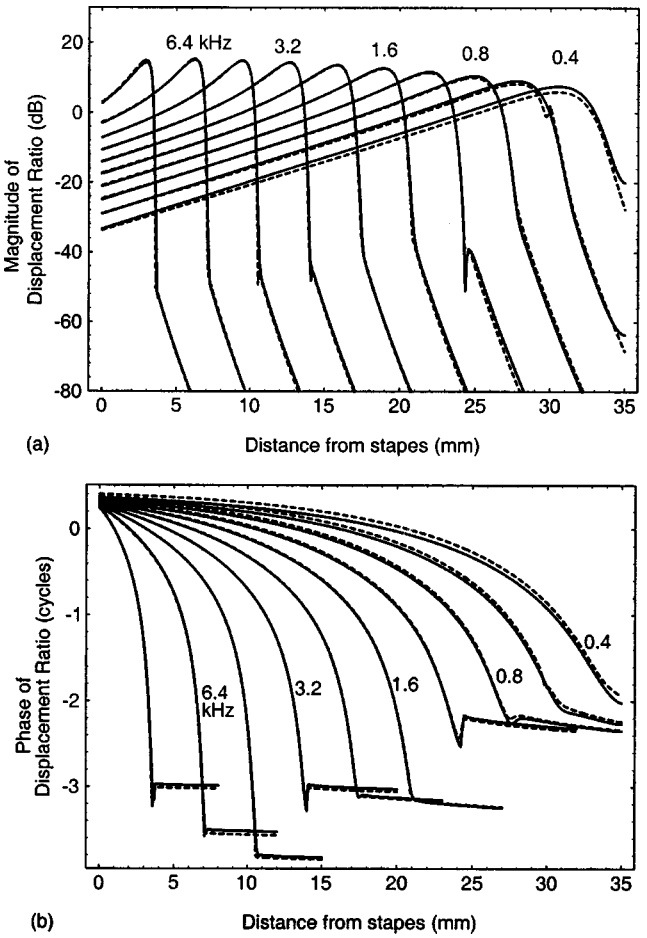


FIG. 5. Comparison of mode-coupling Liouville–Green (dashed) and finite-difference (solid) methods.

This equation has the form

$$c''(x)+P(x)c'(x)+Q(x)c(x)=R(x), \tag{12}$$

that is, it is a second-order ordinary differential equation in $c(x)$, with nonconstant coefficients given by

$$P(x)=\frac{2\int_0^h(\partial\phi_2/\partial x)\,dy}{\int_0^h\phi_2 dy}, \tag{13}$$

$$Q(x)=\frac{\int_0^h\nabla^2\phi_2 dy}{\int_0^h\phi_2 dy}, \tag{14}$$

$$R(x)=-\frac{\int_0^h\nabla^2\phi_1 dy}{\int_0^h\phi_2 dy}. \tag{15}$$

It is possible to obtain approximate closed-form expressions for the preceding integrals; the lengthy formulas are given by Watts (1992). The boundary conditions for the problem are $c(x)=0$ at $x=0$ and $c'(x)=0$ at $x=L$. Thus, we have a one-dimensional boundary-value problem in $c(x)$ with nonconstant coefficients, which may then be solved numerically for $c(x)$.

V. RESULTS AND DISCUSSION

The mode-coupling Liouville–Green (MCLG) approximation is shown in Fig. 5. The MCLG approximation shows good agreement with the numerical solution in both magnitude and phase of the displacement ratio responses for all frequencies. Clearly, the procedure has predicted the correct amount of energy to couple into the k_2 solution, resulting in good agreement with the numerical solution.

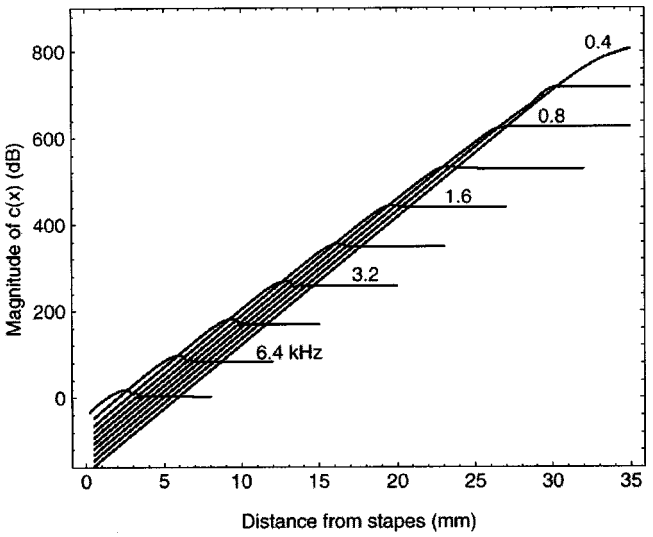


FIG. 6. Magnitude of coupling coefficients $c(x)$ in dB for Neely's parameters. Input frequencies are shown in kHz.

The magnitude of the corresponding coupling coefficients $c(x)$ is shown in Fig. 6. Although the coupling coefficients increase dramatically, they are primarily balancing the natural decay of the cutoff mode. Closer examination of the solutions by Watts (1992) shows that the cutoff mode makes a negligible contribution to the amplitude of the solution at the basilar membrane until the traveling wave mode begins to decay sharply. At this point, the cutoff mode has accumulated significant energy, which it dissipates gradually after the response peak.

In order for the solution to truly satisfy Laplace's equation, it would require contributions from all of the wave modes, not just the traveling-wave and cutoff modes. However, the other modes, by definition, are more heavily damped than the traveling-wave and cutoff modes in their respective regions of dominance, and thus they have only a small local effect which decays quickly after the best place.

The precise relative phase of the two modes in the MCLG approximation is a sensitive function of the physical parameters and the input frequency, and in general may take on any value. Occasionally, the two modes may be exactly out of phase at the basilar membrane position where their amplitudes are equal. In such a case, destructive interference will occur, resulting in a noticeable notch in the amplitude response in the MCLG approximation, as seen, for example, in Fig. 5(a) in the 4.53-kHz trace (third dotted line from the left). In this particular example, the MCLG approximation shows a notch while the numerical solution at that frequency does not, whereas at 1131 Hz [fourth curve from the right in Fig. 5(a)] the numerical solution shows a large notch while the MCLG approximation does not. Clearly, then, both solutions are capable of producing a notch, but they appear to disagree as to when a notch will occur. Since the MCLG solution is only approximate, small errors in the relative phase of the traveling-wave and cutoff modes may cause the

notch to appear or disappear in disagreement with the numerical solution. So, the MCLG approximation provides a mechanism and interpretation of the notch (destructive interference of wave modes), but not a reliable prediction of when it will occur, due to a high sensitivity to errors in the relative phases of the modes.

The above argument was made by Watts (1992), showing that the notch in the MCLG approximation at 4.53 kHz was caused by a 180-deg phase difference between the traveling-wave and cutoff modes, but no claim was made about a notch appearing in the numerical solution at that frequency. Parthasarathi, Grosh, and Nutall (2000) verified that indeed, no notch appears in the numerical solution at that frequency if adequate spatial sampling is used, although they showed that a notch can be artificially introduced in the numerical solution by insufficient spatial sampling. They concluded, however, that the notch in the numerical solution in a two-dimensional model is purely an artifact of insufficient spatial sampling, and thus that two-dimensional models are incapable of demonstrating the notch effect. We verified that real notches do occur in numerical solutions of the two-dimensional model [in agreement with Steele and Taber (1979) and Neely (1981)], for example in Fig. 5(a) at 1131 Hz, and thus we assert that there is no such limitation in the two-dimensional formulation.

A similar notch and change in slope was first observed by Rhode in 1971 in squirrel monkey basilar membrane responses. Rhode concluded: ''It is possible that there is another mode of vibration present in the cochlea.'' We observe that the mode-coupling Liouville–Green approximation predicts the notch and change of slope in agreement with numerical solutions, and with the same qualitative behavior as observed in biological cochleas.

However, it must be noted that the present model neglects the fast acoustic wave associated with the slight compressibility of the real cochlear fluid. Cooper and Rhode (1996) found a notch in the amplitude response above the best frequency, and concluded that it was the result of an interaction between the fast acoustic wave and the normal (nonacoustic) traveling wave mode. This finding, taken together with the present work, suggests that a notch observed in biological measurements above the best frequency could have at least two causes: an interaction of the traveling wave with the fast acoustic wave, or an interaction of the traveling wave with the cutoff mode described in this paper.

VI. CONCLUSIONS

The Liouville–Green approximation to the two-dimensional cochlear mechanics problem fails to solve Laplace's equation in the cochlear fluid. A new solution, called the mode-coupling Liouville–Green approximation, is proposed in which energy is coupled from the traveling wave mode into a second cutoff wave mode. The new approximation gives good agreement with the finite-difference solution. The second wave mode proposed in this paper may provide an explanation for the notch and change in slope of the basilar membrane responses first observed by Rhode in 1971.

ACKNOWLEDGMENTS

The author would like to thank Egbert de Boer, Carver Mead, Dick Lyon, Christopher Shera, Gerald Whitham, Ellen Randall, John Cortese, and Allan Crawford for encouragement and helpful discussions on this work.

APPENDIX: THE CONVENTIONAL VALIDITY CRITERION AND ITS RELATIONSHIP TO THE RLE CRITERION

For a simple long-wave model (which applies for the cochlea near the stapes), the system is governed by the equation

$$\frac{\partial^2 \phi(x,\omega,t)}{\partial x^2} = -k^2(x)\phi(x,\omega,t), \tag{A1}$$

where

$$k(x) = \sqrt{\frac{2\rho\omega^2}{h(S(x)+i\beta(x)\omega - M\omega^2)}}. \tag{A2}$$

In the classic development (Bender and Orszag, 1978) we assume a solution of the form

$$\phi(x,\omega,t) = \phi_0 a(x) \exp i(b(x)+\omega t). \tag{A3}$$

Substituting into Eq. (A1) yields

$$k^2 - b'^2 + 2ia'b'/a + ib'' + a''/a = 0. \tag{A4}$$

By setting the first two terms equal, we get the *eikonal* equation

$$b(x) = \int_0^x k(x)dx, \tag{A5}$$

and by setting the second two terms equal, we get the *transport* equation

$$a(x) = k^{-1/2}(x). \tag{A6}$$

The Liouville–Green approximation amounts to neglecting the a''/a term. The Liouville–Green approximation is generally considered valid (Viergever, 1980; Zweig *et al.*, 1976) when the order of magnitude of the terms in the eikonal equation is much larger than the order of magnitude of terms in the transport equation, leading to the ''conventional validity criterion''

$$\left|\frac{k'(x)}{k(x)^2}\right| \ll 1, \tag{A7}$$

which states loosely that the wavelength should not change too fast on the scale of a single wavelength for the LG solution to be valid. This criterion becomes large near the stapes for low frequencies, and has been used to explain the poor agreement at low frequencies between the LG approximation and the numerical solution (Steele and Taber, 1982). In fact, the poor match at low frequencies is the result of using a first-order stapes displacement term, rather than the higher-order stapes displacement term of Eqs. (4) and (5) (Watts, 1992). When the higher-order term is used, the match at low frequencies is very good, as seen in Fig. 2. So, this leaves a conundrum, namely, if the LG approximation is good at low

frequencies, why does the conventional validity criterion say that the LG approximation is invalid?

The reason is that the conventional validity criterion is only correct to first order. The functions $a(x)$ and $b(x)$ are designed to exactly cancel the first two pairs of terms in Eq. (A4), so the validity of the approximation cannot depend on the relative magnitudes of those terms (since they have been canceled away), but depends on the magnitude of the a''/a term relative to the largest term present, which is k^2; thus, a second-order-correct validity criterion is

$$\left|\frac{a''(x)}{a(x)k^2(x)}\right| \ll 1, \tag{A8}$$

which is the relative error introduced into Eq. (A1) by using the LG approximation. Under the scaling assumptions of Table I, in the long-wave region at low frequencies near the stapes, Eq. (A8) reduces to

$$\left|\frac{1}{16d^2k^2(x)}\right| \ll 1. \tag{A9}$$

In terms of wavelength $\lambda = 2\pi/k$, Eq. (A9) reduces to

$$\lambda \ll 8\pi d, \tag{A10}$$

which is satisfied for all the curves in Fig. 2 prior to the response peak, while Eq. (A7) is not. Note that Eq. (A9) does blow up in the limit as $k \to 0$, just much more slowly than the conventional validity criterion of Eq. (A7).

In the two-dimensional case, it can be shown that the relative error introduced into the governing equation (Laplace's equation) by the approximate solution of Eqs. (1) and (2) is given by the second-order Laplace error (SOLE)

$$\left|\frac{k'}{k^2}\left(1-\frac{4kh(1-kh\tanh kh)}{2kh+\sinh 2kh}+2k(y\tanh ky-h\tanh kh)\right)+\frac{a''}{ak^2}\right| \ll 1, \tag{A11}$$

that is, the relative Laplace error term (RLE) of Eq. (8) plus the second-order relative error term of Eq. (A8).

Since the reasoning behind Eqs. (A8)–(A10) leads to the conclusion that the second-order relative error term of Eq. (A8) is small, even near the stapes at low frequencies, we can drop it, leaving the two-dimensional relative Laplace error of Eq. (8) as the dominant term in the error in the governing equation

$$\left|\frac{k'}{k^2}\left(1-\frac{4kh(1-kh\tanh kh)}{2kh+\sinh 2kh}+2k(y\tanh ky-h\tanh kh)\right)\right| \ll 1. \tag{A12}$$

Now, finally, it can be seen that RLE$\to 0$ as $k \to 0$. This disagrees with the conventional validity criterion, which blows up quickly as $k \to 0$, but the problem lies with the conventional validity criterion, since it is not correct to second order, and does not actually measure the relative error in the governing equation. The analysis behind Eqs. (A8)–(A11) indicates that second-order error terms are small for reasonable choices of parameters such as Neely's. Thus, we conclude that the relative Laplace error accurately indicates the error in the governing equation, and that the conventional validity criterion incorrectly warns of a failure in the LG approximation near the stapes when the approximation is, in fact, a good one in that region.

Bender, C.M., and Orszag, S.A. (**1978**). *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York).

Cooper, N.P., and Rhode, W.S. (**1996**). ''Fast travelling waves, slow travelling waves and their interactions in experimental studies of apical cochlear mechanics,'' Aud. Neurosci. **2**, 289–299.

de Boer, E., and Viergever, M.A. (**1982**). ''Validity of the Liouville–Green (or WKB) method for cochlear mechanics,'' Hear. Res. **8**, 131–155.

Neely, S.T. (**1981**). ''Finite difference solution of a two-dimensional mathematical model of the cochlea,'' J. Acoust. Soc. Am. **69**, 1386–1393.

Parthasarathi, A.A., Grosh, K. and Nuttall, A.L. (**2000**). ''Three-dimensional numerical modeling for global cochlear dynamics,'' J. Acoust. Soc. Am. **107**, 474–485.

Rhode, W.S. (**1971**). ''Observations of the vibration of the basilar membrane in squirrel monkeys using the Mossbauer technique,'' J. Acoust. Soc. Am. **49**, 1218–1231.

Steele, C.R., and Taber, L.A. (**1979**). ''Comparison of WKB and finite difference calculations for a two-dimensional cochlear model,'' J. Acoust. Soc. Am. **65**, 1001–1006.

Viergever, M.A. (**1980**). *Mechanics of the Inner Ear: A Mathematical Approach*, Doctoral dissertation (Delft University Press, Delft, The Netherlands, 1980).

Watts, L. (**1992**). ''Cochlear Mechanics: Analysis and Analog VLSI,'' Doctoral dissertation, California Institute of Technology, 1992. Thesis and MATHEMATICA code are available on-line at http://www.lloydwatts.com/thesis.html

Zweig, G., Lipes, R., and Pierce, J.R. (**1976**). ''The cochlear compromise,'' J. Acoust. Soc. Am. **59**, 975–982.

Noise-evoked otoacoustic emissions in humans

Bert Maat[a)]
Clinical and Experimental Audiology, Academic Medical Center, P.O. Box 22660, 1100 DD Amsterdam, The Netherlands

Hero P. Wit
ENT Department, University Hospital Groningen, P.O. Box 30.001, 9700 RB Groningen, The Netherlands

Pim van Dijk
Department of Otolaryngology and Head & Neck Surgery, University Hospital Maastricht, P.O. Box 5800, 6202 AZ Maastricht, The Netherlands

(Received 16 February 2000; revised 27 July 2000; accepted 28 July 2000)

Click-evoked otoacoustic emissions (CEOAEs) and acoustical responses evoked by bandlimited Gaussian noise (noise-evoked otoacoustic emissions; NEOAEs) were measured in three normal-hearing subjects. For the NEOAEs the first- and second-order Wiener kernel and polynomial correlation functions up to the sixth order were calculated by cross correlating the noise stimulus and the emission response. Emission components were only found in the first-order Wiener kernel. These kernels were dependent on the stimulus level, indicative of the nonlinear response of the cochlea. For each subject the frequency contents of the CEOAE and the first-order Wiener kernel were nearly identical. In the second-order Wiener kernels and higher-order polynomial correlation functions no significant contributions were identified. Apparently, odd and even order nonlinearities cannot be detected by Wiener kernel analysis. © *2000 Acoustical Society of America.* [S0001-4966(00)00911-5]

PACS numbers: 43.64.Jb, 43.64.Kc [BLM]

I. INTRODUCTION

The human cochlea is a sensitive sound transducer, which is capable of detecting sounds in a wide range of intensities and with a high frequency selectivity. This high sensitivity of the cochlea cannot easily be explained by a passive mechanism. Gold (1948) postulated a theory stating that to account for the high frequency selectivity the filters in the cochlea must be active, thus using a feedback mechanism. This mechanism will sharpen up the auditory filters, but can also oscillate if too much energy is fed into them. Gold predicted that oscillating auditory filters may provide acoustical vibrations in the inner ear, which might be transmitted into the external ear canal. About 30 years later, Kemp (1979) discovered that the ear indeed can emit sounds, which can be measured in the ear canal. The presence of these ''otoacoustic'' emissions was a prediction of Gold's theory.

Since 1979, variations of otoacoustic emissions (OAEs) elicited by different kinds of acoustical stimuli have been measured and thoroughly described. It has become clear that OAEs are a phenomenon related to (nearly) normal hearing, generally absent in ears with a light to moderate sensory-neural hearing loss (Probst *et al.*, 1991). Obviously, the emission-generating mechanism is related to normal functioning of the cochlea.

Despite the fact that emission phenomena have been studied extensively during the past 20 years, the emission-generating mechanism in the cochlea is still not resolved. However, it is clear that nonlinear mechanisms are involved in the generation of OAEs.

A general approach to system identification is Wiener kernel analysis. In Wiener kernel analysis, the system under study is described by a set of Wiener kernels (see below). A principal advantage of Wiener kernel analysis is its nonparametric nature, e.g., the behavior of the system can be described without any *a priori* knowledge of its components. On the other hand, with some additional knowledge of the system components, Wiener kernel analysis may be used to identify the component parameters (Van Dijk *et al.*, 1997b). For these reasons, we investigated the application of Wiener kernel analysis to otoacoustic emissions from the inner ear. A preliminary report was published by Maat *et al.* (1994) and comparable results obtained from one frog were described by Van Dijk *et al.* (1997a).

II. WIENER AND VOLTERRA KERNELS

Before introducing Wiener kernel analysis, we need to describe the related concept of Volterra kernels. As Volterra showed, the output $y(t)$ of a nonlinear system S, in response to a time-dependent stimulus $x(t)$ can be written as a series of independent subsystems $H_n[x(t)]$ (Marmarelis and Marmarelis, 1978)

$$y(t)=\sum_{n=0}^{\infty} H_n[x(t)]. \tag{1}$$

Here,

a)Electronic mail: A.Maat@amc.uva.nl

$$H_n[x(t)] = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} h_n(\tau_1,\ldots,\tau_n) x(t-\tau_1),\ldots,$$

$$x(t-\tau_n)\,d\tau_1,\ldots,d\tau_n$$

is the nth order Volterra integral, including the nth order Volterra kernel $h_n(\tau_1,\ldots,\tau_n)$. The number of subsystems, H_n, needed to describe the relation between the system input $x(t)$ and its output $y(t)$ depends on the orders of nonlinearity in the unknown system. For instance, a system consisting of a quadratic nonlinearity in combination with one or more linear filters can be fully described with the first- and second-order kernels $h_1(\tau_1)$ and $h_2(\tau_1,\tau_2)$ (Schetzen, 1989).

Another instructive example is a static nonlinear system [e.g., with y some polynomial of x, or $y=\exp(x)$]. Then, the Volterra kernels h_n are δ functions, and the Volterra series becomes a Taylor expansion.

In practice Volterra kernels are hard to determine. However, when the Volterra functionals are rearranged in linear combinations of H_n, H_{n-2} etc., analogous to the rearrangement of Taylor polynomials into Hermite polynomials, the so-called Wiener G-functionals are formed (Schetzen, 1989)

$$y(t) = \sum_{n=0}^{\infty} G_n[k_n; x(t)]. \tag{2}$$

Like the Volterra expansion in Eq. (1), the Wiener kernel expansion decomposes the system under study into a series of subsystems $G_n[k_n;x(t)]$. Unlike the Volterra subsystems, the Wiener G-functionals can be obtained without *a priori* knowledge of the system: For a Gaussian white-noise stimulus signal, $x(\tau)$, the G-functionals are orthogonal with respect to this stimulus (Schetzen, 1989). This implies that the Wiener kernels can be obtained independently, without knowing the highest order of nonlinearity of the system. When using a Gaussian noise stimulus, the Wiener kernels k_n are given by a set of cross-correlation functions

$$k_0 = \overline{y(t)},$$

$$k_1(\tau_1) = \frac{1}{A}\overline{y(t)x(t-\tau_1)}, \tag{3}$$

$$k_2(\tau_1,\tau_2) = \frac{1}{2A^2}\overline{y(t)x(t-\tau_1)x(t-\tau_2)} - \frac{1}{2A}k_0\delta(\tau_1-\tau_2),$$

where

$$\overline{y(t)} = \lim_{T\to\infty}\frac{1}{T}\int_0^T y(t)\,dt$$

stands for the ensemble average of the time signal and A is the power spectral density of the Gaussian white noise.

Calculation of kernels is manageable up to the third order. Higher-order nonlinearities can be accessed in an economical way with the polynomial correlation method (De Boer, 1979; Eggermont, 1993)

$$p_n(\tau) = \frac{1}{n!}\overline{y(t)\mathrm{He}_n[x(t-\tau)]}, \tag{4}$$

where

$$\mathrm{He}_n[x] = x^n - \binom{n}{2}\frac{A}{2}x^{n-2} + 3\binom{n}{4}\left(\frac{A}{2}\right)^2 x^{n-4} - \cdots,$$

are Hermite polynomials. It can readily be seen that $p_n(\tau)$ is proportional to the diagonal $k_n(\tau,\tau,\ldots,\tau)$ of the nth order Wiener kernel. Although the diagonals do not provide all the information present in the full kernels, for many systems the off-diagonal components of the various kernels are related (see for example the sandwich model as described in Van Dijk *et al.*, 1997b). Thus, for practical system identification, computation of the diagonals only is often sufficient, and leads to a useful description of the system (Eggermont *et al.*, 1983; Van Dijk *et al.*, 1997a).

Calculation of cross-correlation functions for responses evoked by acoustic noise stimulation has been applied in auditory research in brainstem-evoked response audiometry (Dobie and Clopton, 1980; Dobie and Wilson, 1984) and on neural responses with the reversed correlation technique (De Boer, 1967, 1968; Eggermont *et al.*, 1983). Wiener kernel analysis has been performed on anteroventral cochlear nucleus neurons (Wickesberg *et al.*, 1984) and on nerve-fiber responses of bullfrogs (Van Dijk *et al.*, 1994).

III. MATERIAL AND METHODS

In three human subjects with normal hearing, two consecutive emission measurements were performed. A subject was placed in a soundproof booth and a custom-built coupler was connected to the subject's ear canal. This coupler consisted of a miniature electrical driver and a Sennheiser KE 13-227 microphone. The stimulus was generated by a Stanford PG535 pulse generator (100-μs-wide rectangular electrical pulse, with a repetition rate of 8 Hz) or delivered via a Sony PCM-F1 recorder (bandlimited white Gaussian noise, 0–5 kHz). The microphone signal was preamplified with a custom-built circuit, followed by a Philips PM5170 amplifier and bandpass filter (Krone-Hite 3550, 0.3–10 kHz, 24 dB/oct). If click pulses were used as a stimulus, the microphone signal was passed through a custom-built electronic switch (to be described later) before filtering. Both the electric signal delivered to the receiver and the microphone signal were digitally recorded with a two-channel 16-bit Denon DTR-2000 audio DAT recorder, at a sampling frequency of 48 kHz. Recorded signals were transferred off-line to the disk of a NeXT computer for further analysis, using a Singular Solutions A/D 64x.

Stimuli were calibrated in a Zwislocki coupler, using a Brüel & Kjær 4133 measuring microphone connected to a Brüel & Kjær 2636 measuring amplifier. In order to avoid nonlinearities in the receiver and microphone system, noise stimuli were never above 45 dB SPL. The peak level of the click stimulus was 60 peak dB SPL (20 mPa). The long-term power spectral density measured in the Zwislocki coupler was similar for both stimulus signals at maximum levels (see Fig. 1).

For each subject the following measurement series was performed:

At first, we started with six consecutive measurements with a click stimulus. The first measurement was performed at the maximum stimulus level (60 peak dB SPL) for 2 min.

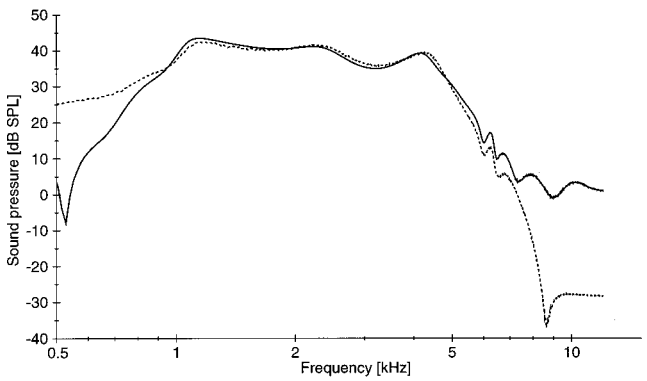


FIG. 1. Amplitude spectrum of the click and noise stimulus at maximum intensity measured in a Zwislocki coupler. The long-term spectrum of the noise stimulus (dashed line) in the coupler is similar to the long-term spectrum of the click stimulus (solid line) in the region of 0.8–5 kHz.

Then, the stimulus was attenuated 10 dB and the measurement was repeated. This was repeated until a stimulus level of 10 peak dB SPL was reached. The two weakest stimulus levels, at 10 and 20 peak dB SPL, were not identifiable by the subjects.

In CEOAE measurements, stimulus and response are both measured in the ear canal. The pressure levels of the click stimuli are three orders of magnitude higher than the response levels. To record both stimulus and response signal in the ear canal, without losing dynamic range of the DAT recorder, the first part of the measured signal (including the stimulus) was compressed with an electronic switch for the first 6 ms after stimulus onset, before being digitally recorded. The compressed signal component was recovered off-line by computer before further analysis. With this technique the dynamic range of the DAT recorder is maximally used for the stimulus and response part in the measured signal. For each stimulus level 1024 artifact-free responses were averaged in the time domain, with a time window of 85 ms.

Second, without removing the probe, six consecutive measurements were performed with bandlimited Gaussian noise as the stimulus, starting at the maximum level of 45 dB SPL and with attenuation steps of 10 dB down to -5 dB-SPL.

Finally, 2 min of recording were performed with no stimulus activity in order to analyze if spontaneous otoacoustic emissions (SOAEs) are involved. SOAEs were calculated by averaging 1024 artifact-free (Hanning) windowed (85 ms) data in the spectral domain using fast Fourier transform (FFT).

For the noise stimuli, both the cochlear response and the stimulus signal are recorded by the recording microphone, e.g.,

$$y(t)=y_{\text{stim}}(t)+y_{\text{coch}}(t)+n(t). \tag{5}$$

Here, the first term y_{stim} represents the acoustic stimulus signal present in the ear canal, due to driving the transducer with an electric signal $x(t)$. The second term $y_{\text{coch}}(t)$ is the noise-evoked otoacoustic response to acoustic stimulus y_{stim}.

The signal $y_{\text{stim}}(t)$ is linearly related to the electrical driver signal $x(t)$. Thus, these two signals are related by a convolution

$$y_{\text{stim}}(t)=[h*x](t), \tag{6}$$

where the system function h described the response of the transducer, the ear canal, and middle ear. For simplicity, we will refer to h as the middle-ear response function.

The third term $n(t)$ is the noise produced by the subject together with the microphone noise, which are both stimulus independent.

The main contribution to the microphone signal corresponds to the linear response $y_{\text{stim}}(t)$. This limits the ability to identify the cochlear response $y_{\text{coch}}(t)$. In order to reduce this problem, we first extracted the cochlear response $y_{\text{coch}}(t)$ out of the full microphone signal $y(t)$. This requires determination of the stimulus signal $y_{\text{stim}}(t)$, which can be computed from the electric stimulus $x(t)$ provided that the middle-ear response function h is known [Eq. (6)]. We estimated this response function by linear cross correlation of the full system response $y(t)$ and the electric signal signal $x(t)$, assuming that the first 3 ms of this cross correlation is the middle-ear response h.

Using the response function h, Eq. (6) was used to compute the signal y_{stim}, and Eq. (5) could then be used to extract the cochlear response y_{coch} from the microphone signal. Next, Wiener kernels were computed by cross correlation of the electrical signal $x(t)$ and the term $[y_{\text{coch}}(t)+n(t)]$, the cochlear response together with the subject and microphone noise.

An illustration of the effect of stimulus subtraction is depicted in Fig. 2, which gives the first 40 ms of the Wiener kernel $k_1(\tau)$ in subject EK. Trace *a* is the calculated $k_1(\tau)$ of the response without subtraction of the stimulus noise; trace *b* is $k_1(\tau)$ with subtraction of the stimulus. For comparison, trace *c* is a CEOAE from the same ear. It can be seen that the signal-to-noise ratio has been improved by using the stimulus subtraction technique. The gained improvement is on average 9 dB.

The unit for the amplitude of click-evoked responses is given in micro Pascal (μPa). It can readily be seen that the unit for the *n*th order Wiener kernel in Eq. (4) is equal to the unit of $y(t)$. The term $y(t)$ corresponds to the measured acoustic signal present in the ear canal [Eq. (5)], which is defined in μPa. So, Wiener kernels can be presented in the same unit as click-evoked responses.

To obtain the diagonals of higher-order kernels we calculated the polynomial correlation functions [Eq. (4)] up to order six. Similar to the calculation of the Wiener kernels, stimulus subtraction was used before calculating these polynomials.

IV. RESULTS

In Fig. 3 the first 4 milliseconds of the acoustic click and the calculated $k_1(\tau)$ for subjects EK, IE, and PE are shown. This represents the response of the probe placed in the ear canal, together with the influence of the middle ear. In each panel, the traces represent responses at the six consecutive attenuation levels used. Traces were superimposed after upscaling to maximum stimulus level. The first 4 milliseconds

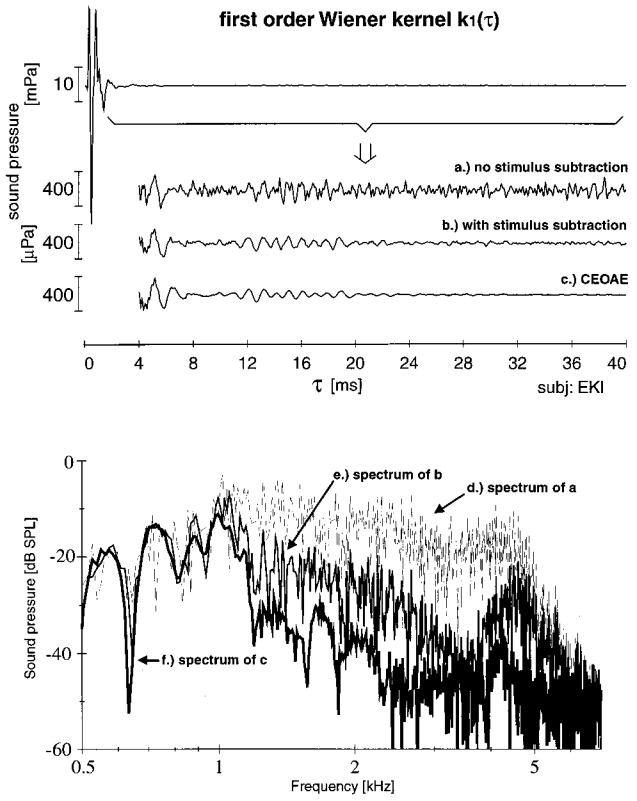


FIG. 2. Upper trace: the calculated first-order Wiener kernel $k_1(\tau)$ for subject EK. Trace *a* is the same kernel without the first 4 milliseconds on a smaller sound-pressure scale. No stimulus subtraction (see the text) was performed before calculating $k_1(\tau)$. Trace *b* represents $k_1(\tau)$ with stimulus subtraction. Trace *c* is a CEOAE from the same ear with the same time and sound-pressure scale. Traces *d*, *e*, *f* are spectra of traces *a*, *b*, *c*.

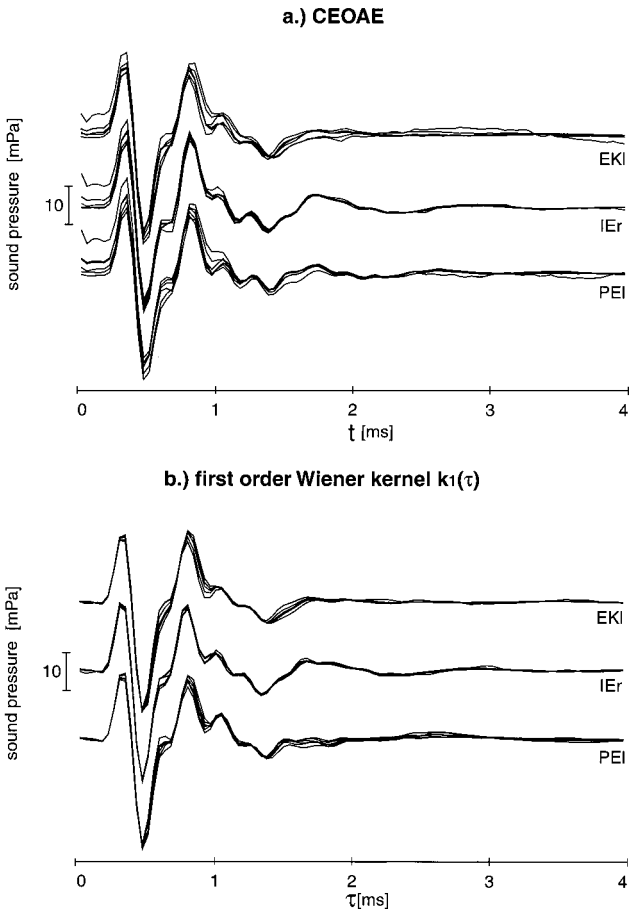


FIG. 3. (a) Six superimposed traces of click-evoked responses for the first 4 ms after click onset for subjects EK, IE, and PE. Stimulus levels are 60 down to 10 peak dB SPL. The traces were multiplied with the stimulus attenuation factor before printing. (b) $k_1(\tau)$ calculated for the same ears for noise stimulus levels of 45 down to -5 dB SPL. No stimulus subtraction was applied before calculating $k_1(\tau)$. The kernels were also multiplied with the stimulus attenuation factor before printing.

of the responses for the different stimulus levels are highly correlated (click: $r=0.98$; noise: $r=0.99$). Also the responses to clicks [panel (a)] and noise [panel (b)] are highly correlated ($r=0.98$). These results confirm linearity of the initial 3 ms of the response. As mentioned in the previous section, the initial 3 ms of the noise response [panel (b)] was used to estimate the middle-ear system function h [Eq. (6)].

In Figs. 4, 5, and 6 CEOAEs and NEOAEs (being the first order Wiener kernels) at different stimulus levels are shown for the three subjects. The first six traces are the CEOAEs at stimulus levels of 60 to 10 peak dB SPL. The next six traces are the calculated first-order Wiener kernels, $k_1(\tau)$, at 45 to -5 dB SPL. For all three subjects, emission responses became detectable at levels between 20 and 30 peak dB SPL for CEOAE measurement and between 15 and 25 dB SPL for the NEOAE measurement. At the two weakest stimulus intensities, both click- and noise-evoked responses diminish into the noise floor.

Subject EK, represented in Fig. 4, had no SOAEs. The subjects IE and PE, represented in Figs. 5 and 6, respectively, both had SOAEs as shown, respectively, in Fig. 7. IE had three emission peaks at 1415, 1511, and 1614 Hz at 5, 32, and 5 dB above noise level. PE had one emission peak at 1936 Hz at 12 dB above noise level.

To detect nonlinear growth of the CEOAE and the first-order Wiener kernel, responses obtained at different stimulus levels were subtracted after proper multiplication to cancel the linear part. In Figs. 8(a) and (b) the result is shown for the four strongest stimulus levels of subject EK; the upper trace of panel (a) and (b) is the response at maximum stimulus level minus $\sqrt{10}$ times the response at 10 dB weaker stimulus level, etc.

For subject EK, only the strongest stimulus levels [upper traces of CEOAE and $k_1(\tau)$, respectively] revealed a significant nonlinear contribution. The corresponding spectra are shown in Fig. 8(c). It is clear that for this subject both the click-evoked and the noise-evoked OAE have a similar content.

Figures 8(d) and (e) show data of subject IE. In contrast to subject EK, IE emitted three SOAEs. The SOAEs clearly contribute to the CEOAE, as can be seen by the prolonged oscillations visible in the CEOAE traces. In contrast, in the noise-evoked OAE SOAEs are suppressed by 10 dB. Nevertheless, spectral analysis still reveals their presence [panel (f)]. The suppressive effect on SOAEs in noise-evoked OAE measurements is also demonstrated in Fig. 9, which shows results for subject IE. Here, the SOAE component, visible in

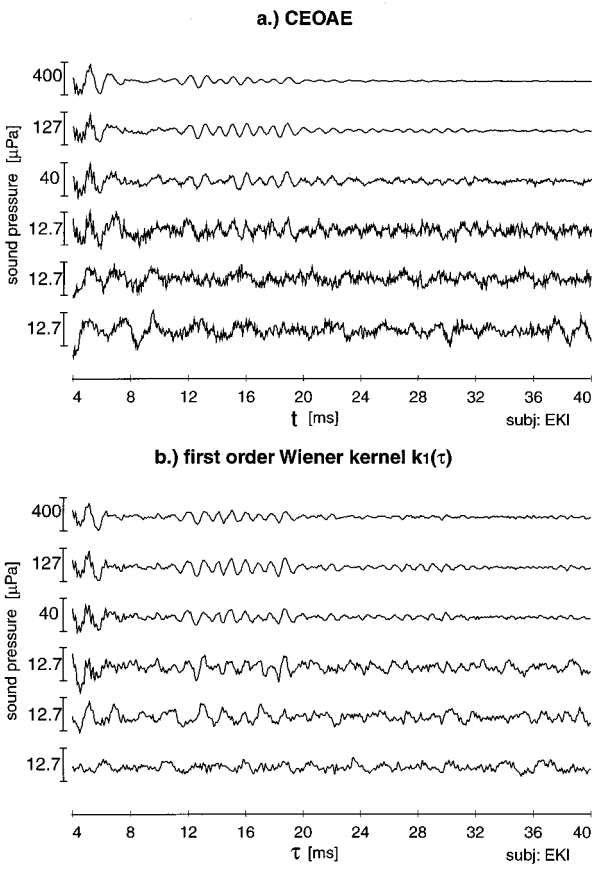


FIG. 4. (a) CEOAE traces from subject EK. Stimulus levels are 60 down to 10 peak dB SPL. The sound-pressure scale for the upper four traces is multiplied with the stimulus attenuation factor in order to make comparison easier. (b) $k_1(\tau)$ calculated for the same ear for noise stimulus levels of 45 down to -5 dB SPL. The same sound-pressure scales as in (a) are used. Stimulus subtraction was applied before calculating $k_1(\tau)$.

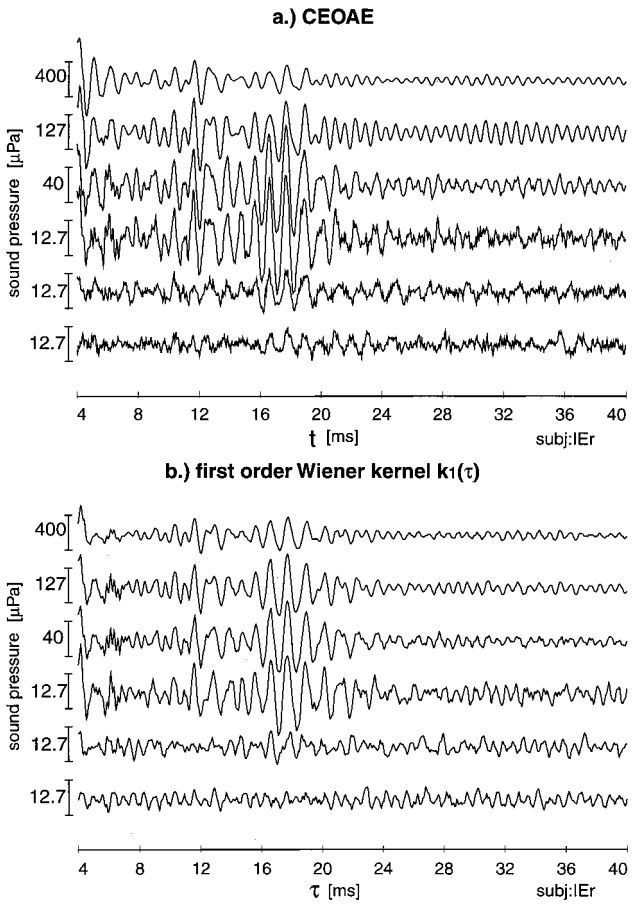


FIG. 5. CEOAE traces and calculated first-order Wiener kernels for subject IE. Stimulus levels and sound-pressure scales were the same as in Fig. 4.

the CEOAE measurement, was suppressed by 6 dB in the NEOAE recording.

To measure the frequency-dependent emission amplitude ratio relative to the stimulus level we computed the envelope of the emissions, using wavelet transformation, with the Morlet wavelet as basis function (Guillemain *et al.*, 1989). Next, for each clearly detectable emission frequency (two per subject) the amplitude growth was calculated. Amplitude growth functions averaged across subjects for both click- and noise-evoked OAEs are depicted in Fig. 10. Both CEOAEs and NEOAEs show a compressive emission amplitude growth, whereas the NEOAE amplitude seems to grow more linearly. This growth difference, however, is not significant ($p<0.05$).

For all subjects the calculation of second-order Wiener kernels did not show identifiable contributions. These kernels are not depicted.

In Fig. 11(a) polynomial correlation functions, obtained from the noise measurement at the maximum stimulus intensity for subject IE, are shown. For all subjects the first-order polynomial correlation function, $p_1(\tau)$, showed emission components. But, $p_2(\tau)$ up to $p_6(\tau)$ showed no contributions (no significant difference compared to correlation functions obtained from measurements in a Zwislocki coupler). Figure 11(b) gives the corresponding spectra of the correlation functions (the first 4 ms were zero-padded before calculating the spectra). Only the Fourier transform of $p_1(\tau)$ (thick line) shows emission peaks.

V. DISCUSSION

A. Correlation technique

Wiener kernel analysis is a nonparametric technique to model stationary stable systems.

Strictly speaking, the technique is not designed for identification of unstable nonlinear systems as, for instance, an ear which produces SOAEs. Nevertheless, we showed that the correlation technique of the Wiener kernel analysis is suitable to describe the response of the ear to a Gaussian noise stimulus.

B. Stimulus subtraction

In click-evoked OAE measurements the first 4 ms of the response, after stimulus onset, consists mainly of the impulse response function of the probe together with ear canal and middle ear. This click artifact persists, despite the use of a nonlinear averaging technique, due to nonlinearities in the measuring system.

A similar problem occurs in Wiener kernel analysis. Then, the first-order Wiener kernel function, i.e., the first-order correlation function of the stimulus noise and the mi-

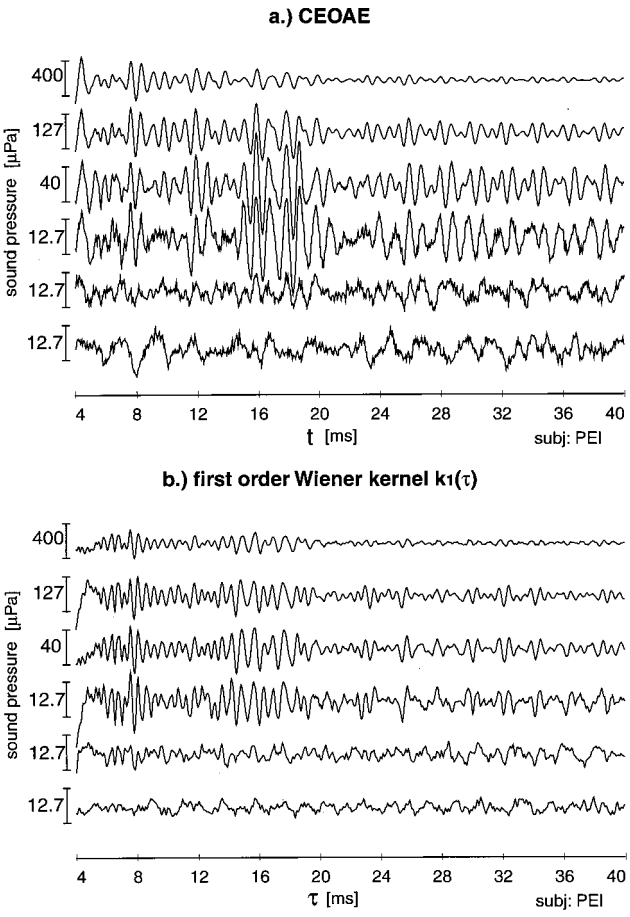


FIG. 6. CEOAE traces and calculated first-order Wiener kernels for subject PE. Stimulus levels and sound-pressure scales were the same as in Fig. 4.

crophone response, contains a large component in the first 4 ms, corresponding to the middle- and outer-ear response. There, the large component is not an artifact of nonlinearities of the measuring system, but is inherent to the linear correlation analysis.

In this work we used the first 4 ms of the Wiener kernel as an estimate of the impulse response function of the outer and middle ear. By combining the recorded stimulus signal and this impulse response function, we estimated the outer- and middle-ear contribution to the recorded ear canal sound pressure. After subtraction of this contribution, we recalculated Wiener kernels, gaining an extra 9 dB signal-to-noise ratio.

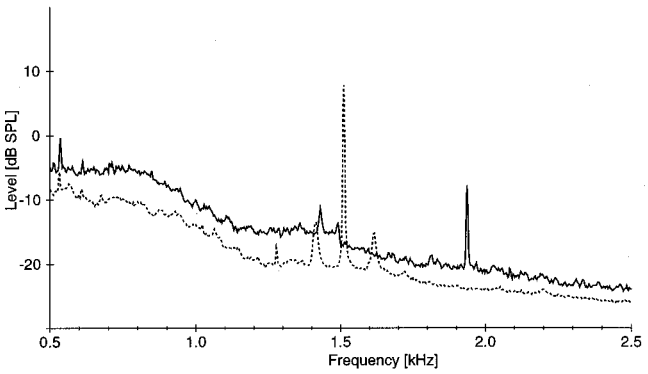


FIG. 7. Power spectra of SOAEs of subject PE (solid line) and IE (dashed line).

C. Wiener kernels

We computed the first- and second-order Wiener kernels, and the higher-order polynomial correlation functions. Only the first-order Wiener kernel displayed emission components. The first-order Wiener kernel was dependent on stimulus level.

The level dependence of the first-order Wiener kernel reflects the nonlinear level-dependent response of the cochlea. In contrast, the absence of emission components in second-order kernel and higher-order polynomial correlation functions suggests a linear cochlear response. Thus, although the response characteristics of the cochlea depend on stimulus level, at each particular noise level it seems to respond linearly. Note that this contrasts the cochlear response to clicks, which seems to consist primarily of intermodulation distortion (Yates and Withnell, 1999).

As mentioned above, the level dependence of the first-order Wiener kernel implies a nonlinear cochlear response. In turn, this implies that the higher-order Wiener kernels must be nonzero. Apparently, the cochlea acts as an automatic gain control (AGC) system, which sets up a (nearly) linear response behavior to each particular noise level. The time constant of the AGC is probably short compared to the measuring time used in this work (120 s), explaining the linear behavior we observed for each given stimulus level. On the other hand, it is long compared to the time base we used for kernel estimation (85 ms). This can be concluded from the fact that no AGC-like contributions are evident in the higher-order kernels we calculated.

D. Comparison of CEOAE-$k_1(\tau)$

The comparison of the first-order Wiener kernels with the CEOAEs at approximately equivalent stimulus levels shows that $k_1(\tau)$ resembles the CEOAE. For the objects, for which Figs. 5 and 6 give the results, this resemblance is good. For subject PE (Fig. 6) a high frequency component is found in the first part of $k_1(\tau)$, which is not present in the CEOAE. The frequency of this component is the same as that for the SOAE of this subject. Apart from these differences, the click- and noise-evoked responses are similar.

The nonlinear residues of $k_1(\tau)$ have the same temporal characteristics as CEOAEs, showing again that there are also nonlinearities in $k_1(\tau)$ (Fig. 8). For the ear without SOAEs [Fig. 8(a)] the comparison is better than in the case with SOAE [Fig. 8(d)]. The weaker the stimulus level the smaller the nonlinear residue.

If no residues are found then, in terms of stimulus level, no nonlinearities are present, as is the case for subject EK in Figs. 8(a) and (b). Residues can also be absent if emissions have a linear behavior with respect to stimulus level. The latter is the case for subject IE, where clear emissions are found down to -40 dB re. maximum stimulus level and where nonlinear contributions [Fig. 8(d)] in the CEOAE case are practically indistinguishable from the noise floor for the

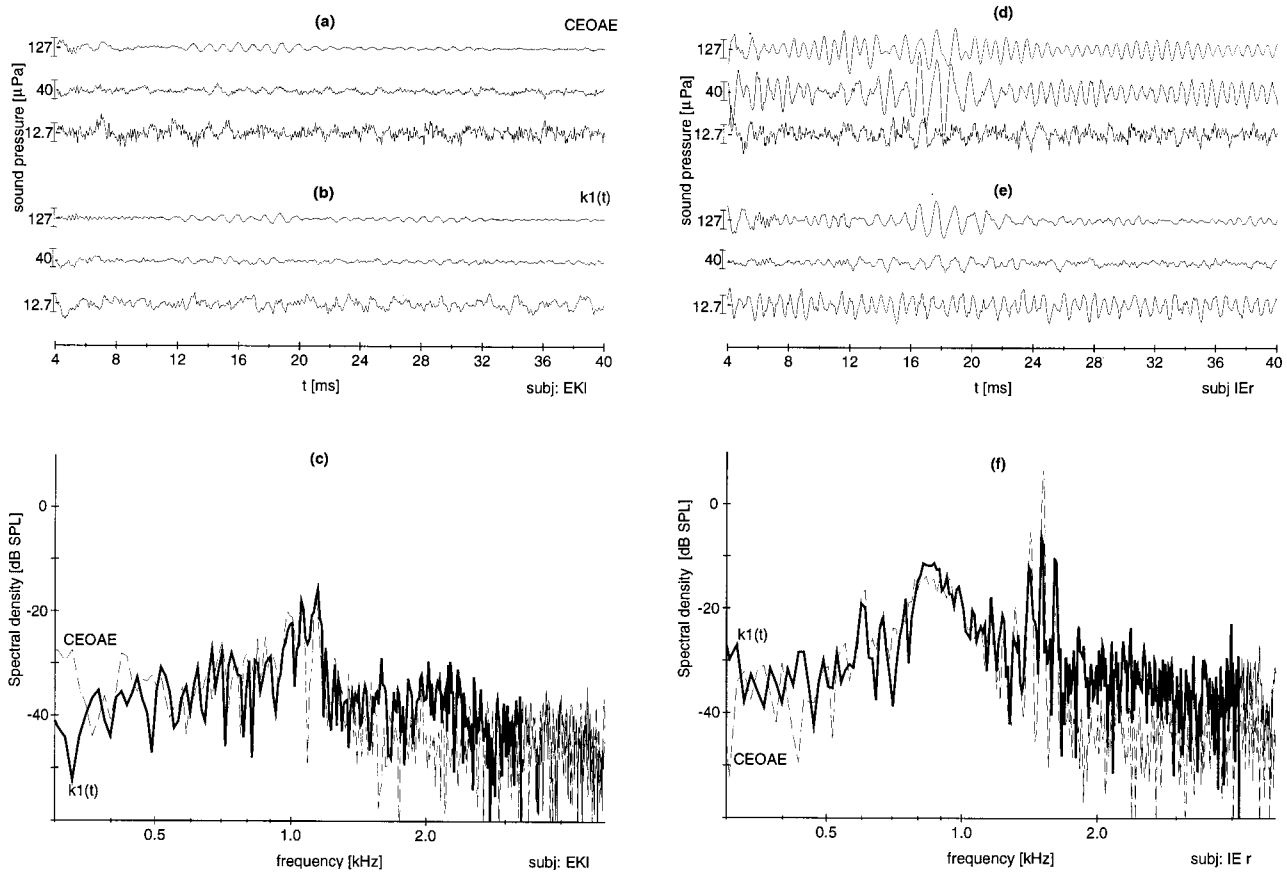


FIG. 8. (a) Nonlinear residue resulting from subtraction of traces in Fig. 4. For CEOAE first trace: [0-dB stim trace] $-\sqrt{10}$[$-$10-dB trace]; second trace: [$-$10-dB stim trace] $-\sqrt{10}$[$-$20-dB trace]; third trace: [$-$20-dB stim trace] $-\sqrt{10}$[$-$30-dB trace]. (b) Analog as for (a) but now for the $k_1(\tau)$ traces. (c) Spectra of the top CEOAE trace (dotted) from (a) and the top $k_1(\tau)$ trace (solid) from (b). (d) Same as (a) but now for the CEOAE traces in Fig. 5. (e) Same as (a) but now for the $k_1(\tau)$ traces in Fig. 5. (f) Same as (c) but now from (d), (e).

lower stimulus levels. For the residues in $k_1(\tau)$ the emission component around 17 ms is canceled out at lower stimulus levels, as in the CEOAE case, but an oscillation remains at the SOAE frequency. Is this oscillatory signal really an emission component that belongs to $k_1(\tau)$?

SOAEs can be synchronized by external tones with frequencies close to the SOAE frequency (Van Dijk and Wit, 1990). Also, a click pulse may synchronize SOAEs (Wit *et al.*, 1981; Moulin *et al.*, 1993). This results in strong emission components in CEOAE measurements at SOAE frequencies, as can be seen in Figs. 5, 6, and 9. The presence of the SOAE component in $k_1(\tau)$ may imply that the SOAE is influenced by, and thus correlated to, the noise. However, a SOAE emission component may also persist in the computed Wiener kernel if it is completely unaffected by the noise: due to the finite measuring time, its presence in the microphone system results in a corresponding component in the kernel.[1]

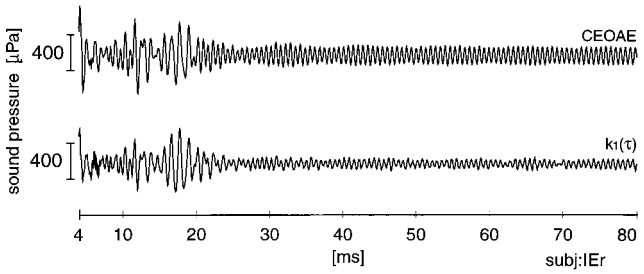


FIG. 9. CEOAE and first-order Wiener kernel $k_1(\tau)$ for subject IE. For the CEOAE the click stimulus level was 50 peak dB SPL and for $k_1(\tau)$ a stimulus level of 35 dB SPL was used.

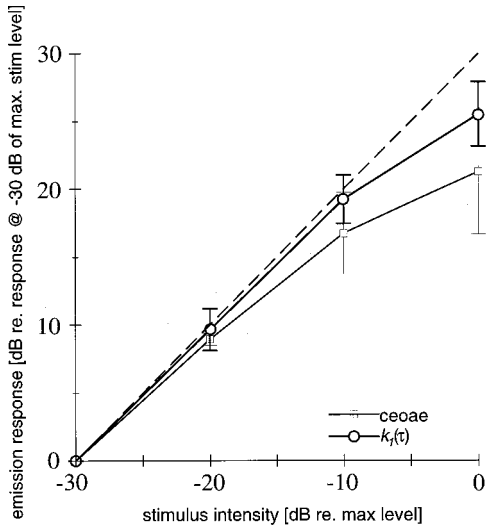


FIG. 10. Emission amplitude ratio with respect to the stimulus intensity (bars: 1 s.d.). The dotted line depicts linear amplitude growth.

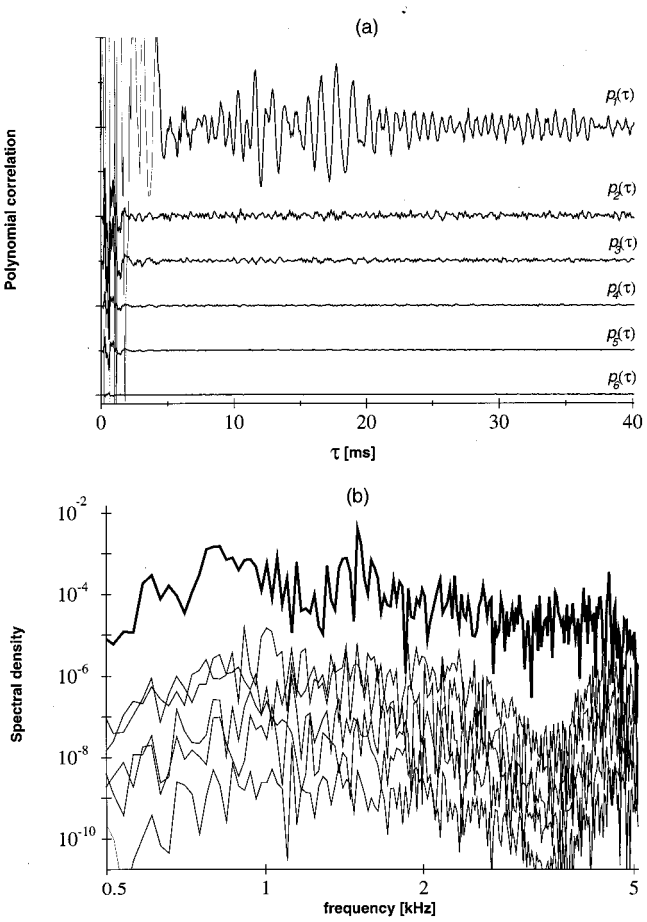


FIG. 11. Polynomial correlation functions obtained from the recordings at maximum stimulus level for subject IE based on the data depicted in Fig. 5. (a) The first-order polynomial correlation function $p_1(\tau)$ up to the correlation function of the sixth order, $p_6(\tau)$. Polynomial correlation functions are proportional to the diagonals of Wiener kernels. (b) Spectra of the polynomial correlation functions.

E. Clinical aspects

Compared to other evoked-emission measurement, a possible advantage of NEOAE measurement for clinical application is that relatively small peak amplitudes are needed in order to obtain a considerable stimulus power. Thus, an emission measurement can be conducted at high stimulus level, without driving the measuring system into saturation. This may be useful when testing patients with mild to moderate hearing losses (conductive or sensory-neural).

F. Conclusion

We demonstrated that OAEs can be found with Wiener kernel analysis using Gaussian white noise as stimulus. The emissions, which were only visible in the first-order Wiener kernel, were consistent with the CEOAE data. The amplitude of $k_1(\tau)$ depended nonlinearly on the stimulus level, similar to the level dependence of CEOAE amplitudes. Apparently, in response to noise the cochlea acts as an automatic gain control system. For such a system, the response is stimulus-level dependent, but at each particular stimulus level it responds linearly.

ACKNOWLEDGMENT

The research of Bert Maat was supported by the Groningen Graduate School for Behavioral and Cognitive Neurosciences (BCN).

[1]If we assume that an ear produces a strong SOAE and the stimulus is too weak to synchronize the SOAE. Then, the SOAE is independent of the stimulus and can, strongly simplified, be described as the function $y_s(t) = B(t)\cos(\omega_0(t+t_0))$. Here, the term $B(t)$ is the amplitude and ω_0 is the frequency of the SOAE. The first-order Wiener kernel calculated of $y_s(t) + y_{\text{other}}(t)$ will then be $k_1(\tau) = B/T\cos(\omega_0\tau + \phi(\omega_0)) + k_{1-\text{other}}(\tau)$. Here, $k_{1-\text{other}}(\tau)$ describes the non-SOAE activity of the ear, T is the averaging time, and $\phi(\omega_0)$ an arbitrary phase. The first part $k_1(\tau)$ is a periodic sinusoidal function on the stimulus noise level. Its amplitude is dependent on the measurement time and will reduce with factor $1/T$. The stimulus noise in $k_1(\tau)$ reduces with factor $1/T^2$, so the sinusoidal oscillation will become visible at kernel-noise level. As strong SOAEs are known as sinusoidal oscillations (Bialek and Wit, 1984) this could explain the existence of the SOAEs in $k_1(\tau)$ at low stimulus levels.

Bialek, W., and Wit, H. P. (**1984**). ''Quantum limits to oscillator stability: theory and experiments on acoustic emissions from the human ear,'' Phys. Lett. **104A**(3), 173–178.

De Boer, E. (**1967**). ''Correlation studies applied to the frequency resolution of the cochlea,'' J. Aud Res. **7**, 209–217.

De Boer, E. (**1968**). ''Reverse correlation. I. A heuristic introduction to the technique of triggered correlation with application to the analysis of compound systems,'' J. Aud Res. **7**, 472–486.

De Boer, E. (**1979**). ''Polynomial correlation,'' Proc. IEEE **67**, 317–318.

Dobie, R. A., and Clopton, B. M. (**1980**). ''Auditory evoked responses obtained by cross correlation: A preliminary report,'' Otolaryngol.-Head Neck Surg. **88**, 797–802.

Dobie, R. A., and Wilson, M. (**1984**). ''Short latency auditory responses obtained by cross correlation,'' J. Acoust. Soc. Am. **76**, 1411–1421.

Eggermont, J. J. (**1993**). ''Wiener and Volterra analysis applied to the auditory system,'' Hear. Res. **66**, 117–201.

Eggermont, J. J., Johannesma, P. I. M., and Aertsen, A. M. H. (**1983**). ''Reverse-correlation methods in auditory research,'' Q. Rev. Biophys. **16**, 341–414.

Gold, T. (**1948**). ''Hearing II. The physical basis of the action of the cochlea,'' Proc. R. Soc. E **B135**, 492–498.

Guillemain, P., Kronland-Martinet, R., and Martens, B. (**1989**). ''Estimation of spectral lines with the help of wavelet transform,'' in *Wavelets and Applications*, edited by Y. Meyer, Research Notes in Applied Mathematics (Springer, Berlin), pp. 38–60.

Kemp, D. T. (**1979**). ''Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea,'' Arch. Otorhinolaryngol. **224**, 37–47.

Maat, B., Van Dijk, P., and Wit, H. P. (**1994**). ''Noise evoked otoacoustic emissions; just an other way of measuring?,'' in *Advances in Otoacoustic Emissions, Volume 2. Recording Techniques for Otoacoustic Emissions*, edited by R. Thornton and F. Grandori (European Concerted Action on Otoacoustic Emissions, Lecco, Italy), pp. 21–31.

Marmarelis, P., and Marmarelis, V. (**1978**). *Analysis of Physiological Systems* (Plenum, New York).

Moulin, A., Collet, L., Veuillet, E., and Morgon, A. (**1993**). ''Interrelations between transiently evoked otoacoustic emissions, spontaneous otoacoustic emissions and acoustic distortion products in normally hearing subjects,'' Hear. Res. **65**, 216–233.

Probst, R., Lonsbury-Martin, B. L., and Martin, G. K. (**1991**). ''A review of otoacoustic emissions,'' J. Acoust. Soc. Am. **89**, 2027–2067.

Schetzen, M. (**1989**). *The Volterra and Wiener Theories of Nonlinear Systems* (Krieger, Malabar, FL).

Van Dijk, P., Maat, A., and Wit, H. P. (**1997a**). ''Wiener kernel analysis of a noise evoked otoacoustic emission,'' Br. J. Audiol. **31**, 473–477.

Van Dijk, P., Wit, H., and Segenhout, J. (**1997b**). ''Dissecting the frog inner ear with Gaussian noise. I. Application of higher-order Wiener-kernel analysis,'' Hear. Res. **114**, 229–242.

Van Dijk, P., and Wit, H. P. (**1990**). ''Synchronization of spontaneous otoacoustic emissions to a $2f_1-f_2$ distortion product,'' J. Acoust. Soc. Am. **88**, 850–856.
Van Dijk, P., Wit, H. P., Segenhout, J. M., and Tubis, A. (**1994**). ''Wiener kernel analysis of inner ear function in the American bullfrog,'' J. Acoust. Soc. Am. **95**, 904–919.
Wickesberg, R., Dickson, J., Gibson, M., and Geisler, C. (**1984**). ''Wiener kernel analysis of responses from the anteroventral cochlear nucleus neurons,'' Hear. Res. **14**, 155–174.
Wit, H., Langevoort, J., and Ritsma, R. (**1981**). ''Frequency spectra of cochlear acoustic emissions ('Kemp-echoes'),'' J. Acoust. Soc. Am. **70**, 437–445.
Yates, G., and Withnell, R. (**1999**). ''The role of intermodulation distortion in transient-evoked otoacoustic emissions,'' Hear. Res. **136**(2), 49–64.

Basilar membrane responses to broadband stimuli

Alberto Recio[a] and William S. Rhode
Department of Physiology, University of Wisconsin, Madison, Wisconsin 53706

(Received 17 May 2000; accepted for publication 23 August 2000)

Basilar membrane (BM) responses to two types of broadband stimuli—clicks and Schroeder-phase complexes—were recorded at several sites at the base of the chinchilla cochlea. Recording sites (characteristic frequency, CF, in the range of 5.5–18 kHz) span the 1–4-mm basal region of the basilar membrane. BM responses to clicks consisted of undamped oscillations with instantaneous frequency that increased over time until it reached a value around CF. The time constant of this glide is CF dependent. Throughout the entire region under study, BM vibration exceeded umbo motion by up to 60 dB. Nonlinear properties of BM responses to clicks resemble those found in the more studied 8–10-kHz region. Amplitude spectra of Schroeder-phase complex stimuli, which consist of a series of sinusoidal components summed in negative (−SCHR) and positive Schroeder phase (+SCHR), are flat. The envelope of BM responses to +SCHR stimuli contains valleys, or dips, that are wider than those found in responses to the −SCHR stimuli. Hence, BM responses to the former stimuli are ''peakier'' than responses to the latter. Differences in response waveforms are less obvious in linear cochleae. Suppression of a near-CF tone by −SCHR stimuli was larger than that evoked by +SCHR stimuli. [S0001-4966(00)05411-4]

PACS numbers: 43.64.Kc [LHC]

I. INTRODUCTION

Although a great deal is known about basilar membrane (BM) responses to tones (e.g., Rhode, 1971; Ruggero *et al.*, 1997; Nuttall and Dolan, 1996), two-tone stimuli (e.g., Ruggero *et al.*, 1992; Rhode and Cooper, 1993; Cooper, 1996; Robles *et al.*, 1997), and to a lesser extent to clicks (Robles *et al.*, 1976; Recio *et al.*, 1998) and noise (de Boer and Nuttall, 1997; Recio *et al.*, 1997), little is known about BM responses to other complex stimuli, such as those consisting of harmonic-related tones or frequency-modulated (FM) sinusoids. Because of the nonlinear nature of BM responses, level-dependent gain functions obtained from steady-state responses to tones cannot necessarily be used to predict responses to clicks or other complex stimuli. This paper examines BM responses to clicks at several locations along the basal end of the chinchilla cochlea as well as responses to Schroeder-phase complexes.

It has been shown that BM responses to clicks and tones generally agree with each other (Robles *et al.*, 1976; Recio *et al.*, 1998), as is always the case for *linear* systems. Robles *et al.* (1976) compared the response to clicks and tones in the time domain, i.e., by using the inverse Fourier transform of the steady-state velocity data (a temporal waveform). Their results show a good match between the ''synthetic'' impulse response and the experimentally obtained one. Recio *et al.* (1998) compared the two sets of responses in the frequency domain, i.e., by using the Fourier transform of the click responses. Again, their results show what was considered a good agreement between the two responses, in spite of the relatively coarse frequency sampling of the tone stimuli (1-kHz increments). Because of the large ringing of some click responses shown recently (Ruggero *et al.*, 1992; Ruggero *et al.*, 1996; Recio *et al.*, 1998), it would be interesting to know if such ringing can be predicted from responses to tones. We therefore decided to compare the inverse transform of gain functions obtained at a frequency sampling of 100 or 250 Hz with responses to clicks. Clicks are also useful in obtaining cochlear delays, and for that purpose we measured click responses at several locations along the BM. Because the hook region in the chinchilla has been little investigated (Narayan and Ruggero, 2000; Rhode and Recio, 2000), we also determined some of the properties of BM responses to clicks in that cochlear region. Click responses from a region more apical than the 8–10 kHz region studied by Ruggero and his colleagues (Robles *et al.*, 1986; Ruggero *et al.*, 1997; Recio *et al.*, 1998) were measured as well.

In addition to BM click responses, we investigated the responses to a harmonic complex sound $m(t)$, referred to in the literature as the Schroeder-phase complex (e.g., Smith *et al.*, 1986; Kohlrausch and Sander, 1995; Carlyon and Datta, 1997; Summers and Leek, 1998),

$$m(t)=\sum_{n=n1}^{n2} A_0 \sin(2\pi n f_0 t+\theta_n).$$

The phase, θ_n, is based on a formula proposed by Schroeder (1970),

$$\theta_n=\pm \pi n(n-1)/N,$$

where $N=n2-n1+1$. For a given fundamental frequency f_0, $m(t)$ can have two wave shapes depending on the sign of the phase θ_n. Stimuli generated with a ''+'' sign will be referred to as +Schroeder-phase complexes (+SCHR, Fig. 1, bottom), whereas −Schroeder-phase complexes (−SCHR, Fig. 1, top) will denote those generated with a ''−'' sign. When each harmonic is summed in phase, $m(t)$ displays a highly peaked waveform. Adding the harmonics with Schroeder phase produces a waveform with a flat envelope. It has been pointed out (Schroeder, 1970) that the instanta-

[a]Electronic mail: recio@physiology.wisc.edu

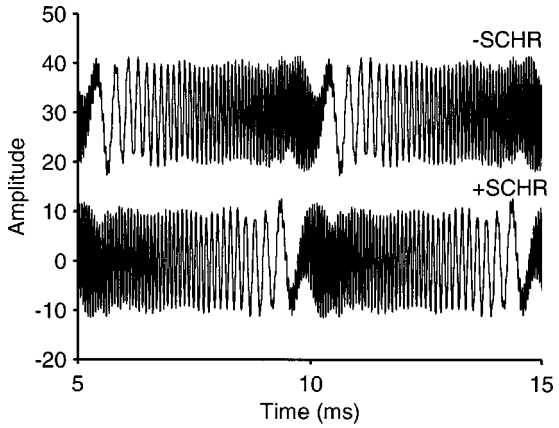


FIG. 1. Waveforms of harmonic complexes for two choices of starting phases. Top waveform: −Schroeder-phase complex (−SCHR), $\theta_n = -\pi n(n-1)/N$. Lower waveform: +Schroeder-phase complex (+SCHR), $\theta_n = +\pi n(n-1)/N$. Total number of components (N) is equal to 96, with harmonics of 5 to 100 of fundamental frequency of 200 Hz. Amplitudes are expressed in arbitrary units.

neous frequency of the Schroeder-phase complex linearly increases (for −SCHR), or decreases (for +SCHR), within each period of f_0. Hence, $m(t)$ can be regarded as either a flat harmonic tone complex or as a series of upward (or downward) frequency sweeps (Smith *et al.*, 1986). These stimuli have been used as maskers in psychoacoustical experiments in humans (e.g., Smith *et al.*, 1986; Kohlrausch and Sander, 1995; Carlyon and Datta, 1997; Summers and Leek, 1998) and it was found in those studies that the masker with the −Schroeder-phase complex always yielded thresholds higher (i.e., greater masking) than those obtained using the +Schroeder-phase complex. Simulation studies (Smith *et al.*, 1986; Kohlrausch and Sander, 1995) using linear transmission line models of BM motion (Strube, 1985) indicated that differences in masking thresholds could be attributed to BM properties. It is not clear, however, what the response of a nonlinear basilar membrane would be, particularly when both masker and signal stimuli are simultaneously presented. Instead of relying on a computer model, we recorded the responses of linear and nonlinear BMs to these stimuli.

II. METHODS

A. Surgical preparation

Twenty-two chinchillas (average weight=500 gm) were anesthetized using pentobarbital sodium (70 mg/Kg, i.p.). Supplementary smaller doses were given as needed to maintain a deeply areflexive state. Seventeen of those chinchillas were used in a recent study of BM responses to tones (Rhode and Recio, 2000). Tracheotomies were performed, but artificial ventilation was used only when necessary. Core body temperature, measured using a rectal probe, was maintained at 38 °C using a thermostatically controlled heating pad. The left pinna was resected, the bulla was widely opened and the stapedius muscle was usually detached from its anchoring. A silver-wire electrode was placed on the round window to record compound action potentials (CAPs) evoked by tone bursts. Gold-coated polystyrene microspheres (25 μm diameter) were introduced into the cochlea in the area called the hook region, after removing the round window membrane, or in a region approximately 2–3 mm away from the hook, after thinning the bone with a microchisel and then removing bone fragments with a metal pick. Except when noted, BM recordings were made after the hole was covered with a piece of coverslip glass. This eliminates most problems associated with movements of the perilymph meniscus (Cooper and Rhode, 1996). The middle-ear response was measured by exposing the umbo through an opening in the external ear canal directly over the tympanic membrane or from the stapes, near the incudostapedial joint. In some experiments, the approximate longitudinal and radial locations of the recording sites were measured *in vivo* with the help of a scale in the microscope's lens. In those measurements, the location of the most basal bead (Fig. 8) is expressed relative to an arbitrary reference.

This protocol was approved by the Animal Care and Use Committee of the University of Wisconsin-Madison.

B. Stimulus generation and control

Acoustic stimuli were generated using an IBM-type computer in conjunction with a Tucker-Davis Technologies system. The acoustic system was a reverse-driven condenser microphone cartridge (Bruel and Kjær model 1434 with squared-root precompensation). Stimuli were delivered to the tympanic membrane through a closed-field sound system. Stimuli used in this project are single tones, condensation clicks (10-μs duration) and Schroeder-phase complexes (20.48-ms duration). The fundamental frequency, f_0, for the Schroeder-phase complexes was always 200 Hz. These stimuli contain all the harmonics between 1000 and 20 000 Hz. When a tone and the Schroeder-phase complex stimuli were played simultaneously, each was presented via a separate microphone. Neither clicks nor Schroeder-phase complexes were compensated by the acoustic calibration, as this was reasonably flat (within 10 dB) in the vicinity of CF. Click levels were expressed as peak pressure *re* 20 μPa, using the procedure described in Recio *et al.* (1998). The level of the harmonic complex was obtained by estimating the ''effective SPL'' in decibels (Kochhar, 1992). Computation of the effective SPL consists of filtering the stimulus by the calibration and estimating the energy at the output of the filter. The sinusoidal stimuli (100–25 000 Hz) were calibrated using a probe tube microphone whose tip was near the tympanic membrane. For the purpose of comparing BM responses to tones and clicks, however, tones were presented at frequencies in the 100–25 000-Hz range with a fixed attenuation level. This was done to ensure that, at least for a linear system, both types of stimuli were equivalent. The number of stimulus presentations was varied (8–1000) in order to maximize the signal-to-noise ratio in the recordings. Stimuli were presented using one or two 16-bit D/As (TDT®) at a playback rate of 200 kHz.

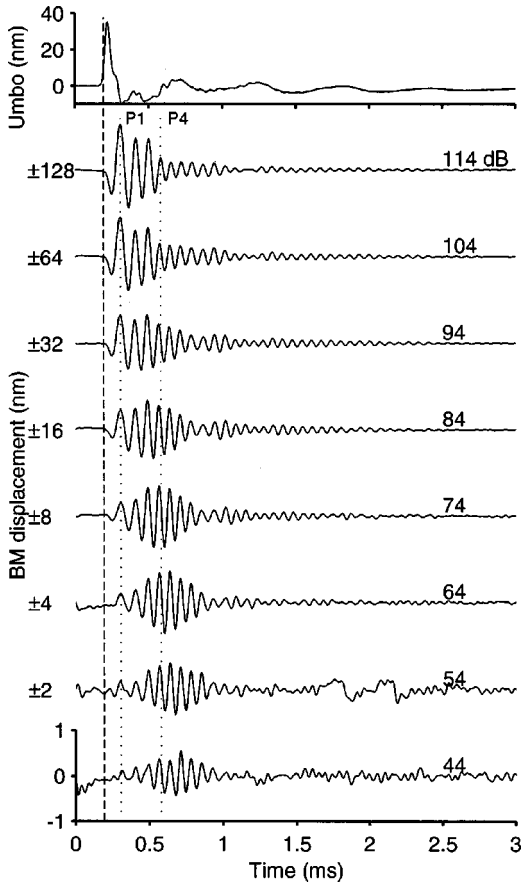


FIG. 2. Basilar membrane (animal CB21, CF=14.5 kHz) responses (lower panel) and umbo response (upper panel) to condensation clicks. Maximum peak intensity of the click stimulus is 114 dB SPL (top trace, lower panel); each of the lower traces displays the response to a click 10 dB less intense than the one above. Positive displacements indicate motion toward scala vestibuli (for BM responses) or toward cochlea (for umbo responses). Number of stimulus presentations is 250. Umbo response was measured in response to a 114-dB SPL click. Number of stimulus presentations is 1000. Dashed line represents umbo onset delay (=192 μs). Onset of BM motion (not shown) was estimated at 220 μs.

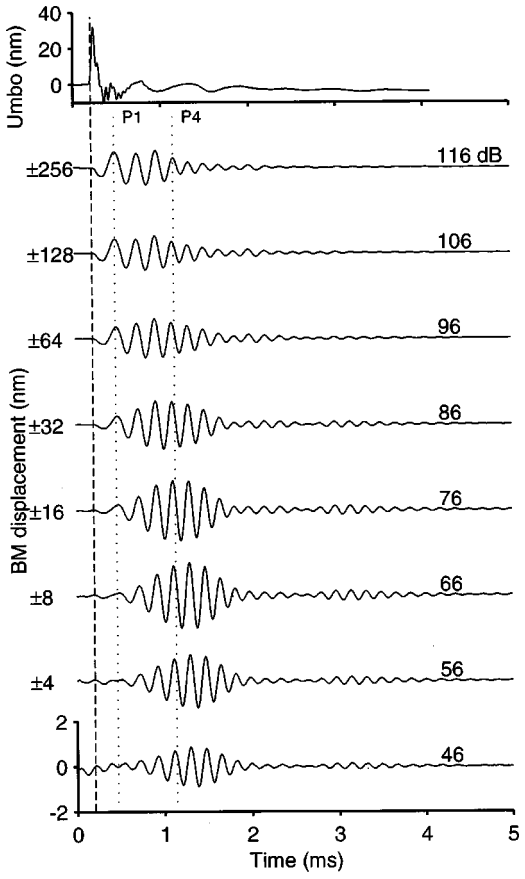


FIG. 3. Basilar membrane (animal CB58, CF=5.5 kHz, lower panel) and umbo responses (upper panel) to condensation clicks. Maximum click intensity is 116 dB SPL (top trace, lower panel); lower traces display the response to a click 10 dB below the one above. Number of stimulus presentations is 5000. Umbo response was measured in response to a 116 dB-SPL click. Number of stimulus presentations is 500. Dashed line represents umbo onset at 204 μs. Onset of BM motion was estimated at 252 μs.

C. Measurement and analysis techniques

Motion in the BM and middle-ear ossicles was recorded using a displacement-sensitive heterodyne laser interferometer (Cooper and Rhode, 1992). The displacement sensitivity of the interferometer is a result of decoding the Doppler phase, as opposed to frequency shifts induced by movements of the reflective targets. An analog phasemeter was used in conjunction with the computer to perform phase unwrapping as necessary. The phasemeter outputs were sampled at a rate of 250 kHz, using a 16-bit A/D (Analogic FAST-16). The noise floor of the recording system was <5 pm/$\sqrt{\text{Hz}}$.

Digitized response waveforms were analyzed using MATLAB® for a sine fitting routine, Fourier analysis or analytic signal representation (Bennet, 1970). The analytic signal of a waveform is a complex quantity whose real part is the waveform itself and whose imaginary part is the Hilbert transform of the real part. The envelope of the waveform is the magnitude of the analytic signal and the instantaneous frequency corresponds to the derivative of the phase of the analytic signal.

III. RESPONSES TO CLICKS

A. Responses to clicks

Figure 2 shows the BM and umbo responses to condensation clicks presented at several intensities. The BM data come from recordings at a site in the hook region (CF=14.5 kHz). BM responses to clicks reflect the characteristic bandpass tuning, while the umbo response (top trace) displays a short oscillation, which is characteristic of a broadly tuned system. Inward displacements of the umbo (i.e., displacements toward the cochlea) are plotted upwards. BM responses are plotted at scales that are 6 dB apart from each other. Positive BM responses indicate displacement toward scala vestibuli. Hence, the first oscillation, or peak, in BM responses to condensation clicks is in the direction of scala tympani. The wave shape of BM responses to clicks is level dependent, as the response peak shifts to earlier times as the level of stimulation increases. Compression in the responses is also evident, particularly at the lowest levels, since the growth in the responses ($\approx$6 dB) is less than the 10 dB difference in stimulus level. Onset delays were defined as the time at which the first peak reached 20% of its maximum absolute value and were estimated for both umbo (Fig. 2, dashed line) and BM motions. Onset delays contain propa-

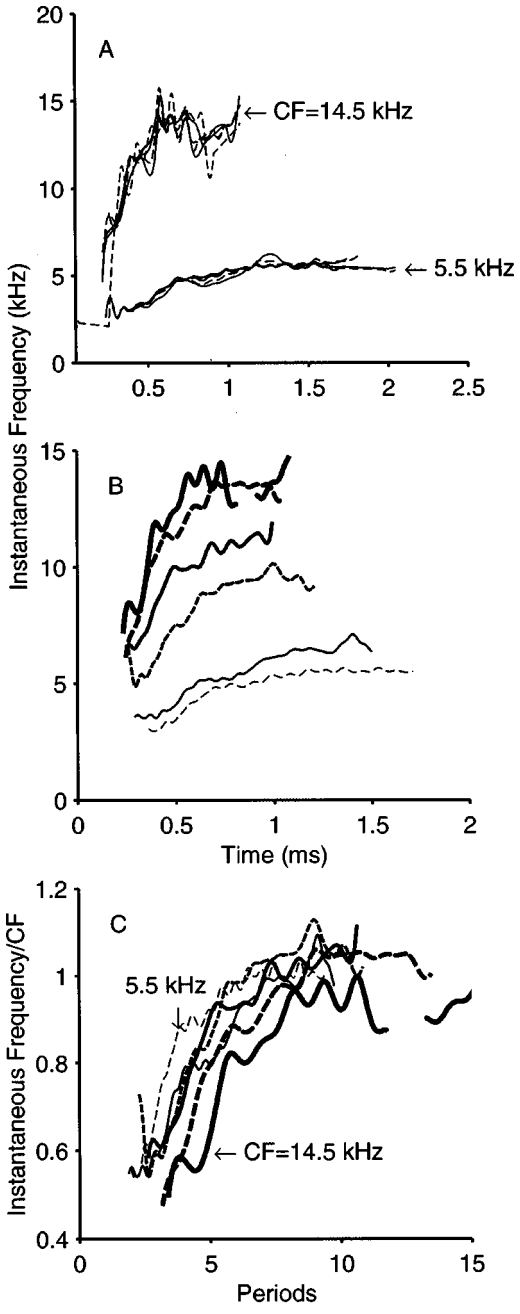


FIG. 4. (a) Instantaneous frequency plotted against time for preparation CB21 (upper curves) and CB58 (lower curves) at several stimulus levels. (b) Instantaneous frequency curves obtained from the responses at six sites in six cochleae. Thickness of the lines is in proportion to CF. (c) Same data as in panel (b), but frequency values normalized to CF and time scale expressed as periods (=1/CF).

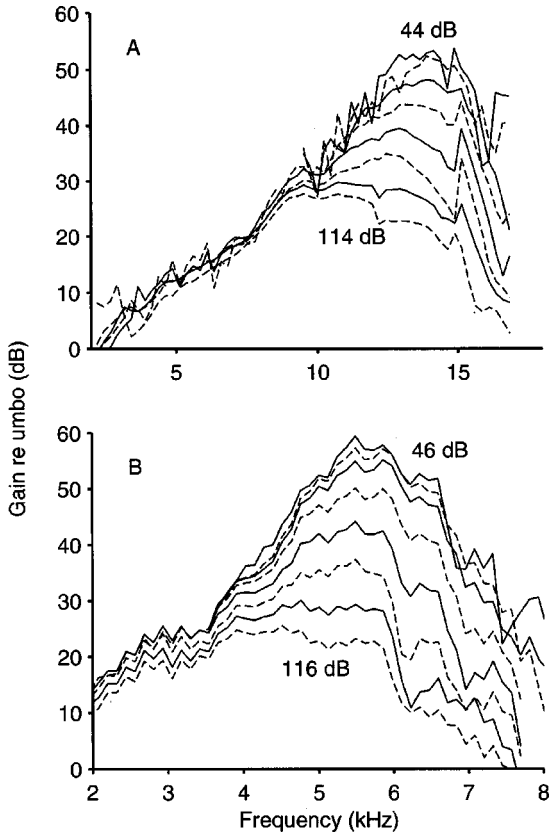


FIG. 5. (a) and (b) Fourier transform magnitudes of basilar membrane click responses as a function of frequency expressed relative to umbo motion for animal CB21 (upper panel) and CB58 (lower panel).

gation delays due to the acoustic system as well as traveling wave times. The cochlear delay (i.e., the difference between BM and umbo onsets) measured at this location was 28 μs. Another characteristic of BM responses to clicks is that the time at which local peaks or zero-crossings occur changes very little, or not at all, with intensity. Dotted lines in Fig. 2 show the times of occurrence of P1 (the first positive peak) and P4 (the fourth positive peak), as measured from the response to the 114-dB click. Whereas the relative position of P1 does not change as a function of level, there are some small changes in the time of occurrence of P4, by as much as 16 μs in waveforms shown in Fig. 2, as it appears to occur at earlier times as the intensity of stimulation decreases.

Figure 3 shows BM and umbo responses to condensation clicks in another preparation (CF=5.5 kHz). Although the difference in CFs is more than 1.5 octaves, responses to clicks in this more apical site display the same nonlinear characteristics described above for the hook region as well as for the 10-kHz region (Recio *et al.*, 1998). The cochlear delay measured at this location was 48 μs. A larger cochlear delay than the one measured in the hook region (Fig. 2) was expected since the recordings come from a more apical site.

B. Frequency modulation in BM responses to clicks

BM responses to clicks show a change in the instantaneous frequency or periodicity as a function of time (deBoer and Nuttal, 1997; Recio *et al.*, 1998). That is, the intervals between successive zero-cross points in the response waveforms are not constant but decrease over time until they reach a steady-state value. The instantaneous frequency of the responses was quantified by means of the analytic signal representation. Figure 4(a) shows instantaneous frequency versus time curves for the data in Figs. 2 and 3. The instantaneous frequency increases in a logarithmic fashion, as determined from analysis not shown, until it reaches a near-CF steady-state value (arrows in Fig. 4). The time it takes to reach the steady state varies as a function of CF, being around 600 μs for the CF=14.5 kHz and 1 ms for the lower-CF data. Although there are some changes in the in-

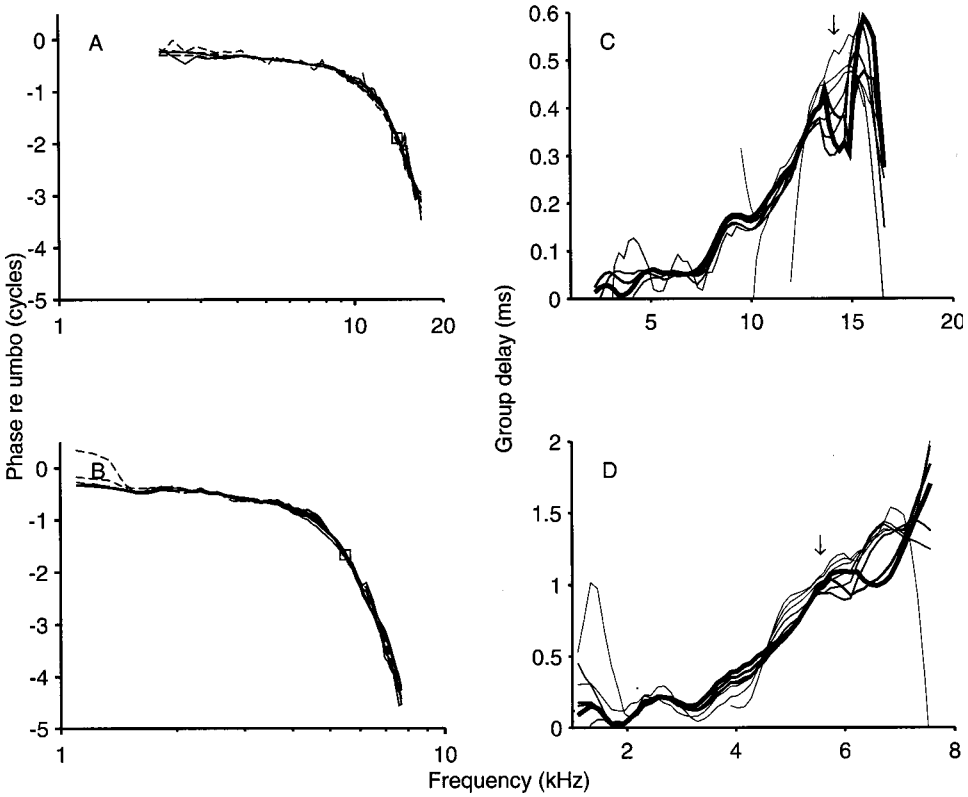


FIG. 6. Basilar membrane phase functions, estimated by Fourier transformation of click responses, at several levels as a function of frequency expressed relative to umbo phase for animal CB21 [panel (a)] and CB58 [panel (b)]. Squares indicate phase at CF. Arrows indicate CF location. Panels (c) and (d) show group delay versus frequency curves at several levels (line thickness is in proportion to stimulus level). Phases in panels (a) and (b) represent phases of BM motion toward scala vestibuli relative to inward movement of umbo.

stantaneous frequency as a function of level, the properties described above remain regardless of the level of stimulation as indicated by the different lines.

Figure 4(b) shows a collection of instantaneous frequency curves obtained from recordings at six sites in several cochleae (5.5 kHz$\leqslant$CF$\leqslant$14.5 kHz). All the curves reach steady-state values approximately equal to the CF of the recording site at a speed in proportion to their CF. Figure 4(c) displays the same curves shown in the middle panel but after normalizing their value to CF. The time scale is also expressed in periods of the characteristic frequency. If the amount of frequency modulation in all the curves were the same, all the curves would overlap each other. The results show that curves originating from lower-CF sites (thinner curves) reach steady state earlier (relative to the number of periods) than those from the highest CFs (thicker curves). This indicates that the relative amount of frequency modulation in the cochlea decreases with increasing CF.

C. Fourier analysis of BM responses to clicks

Figures 5(a) and (b) display Fourier transform amplitudes for the data shown in Figs. 2 and 3 normalized to umbo responses. As expected, BM gain functions around CF increase in value as the intensity of the stimulus decreases, in both hook and more apical regions. The frequency range of the nonlinearity appears to be larger for the preparation with CF=5.5 kHz [Fig. 5(b)] than that measured in the hook region [Fig. 5(a)]. That is, for the 5.5-kHz region, the nonlinearity extends beyond the traditional $\frac{1}{2}$ octave shift (to about 3.5 kHz), whereas in the hook region the frequency-dependent nonlinearity does not extend beyond about 9 kHz (which is approximately $\frac{1}{2}$ octave below CF). In both cases, gain functions show a notch at certain frequencies above CF. This behavior was also observed in previous BM responses to clicks (Recio *et al.*, 1998) and was shown then to be a consequence of phase reversals in the click response. Short-time Fourier analysis performed in Recio *et al.* (1998) showed that phase of a near-CF component during the first millisecond of the response changed by about 180° during later parts of the response.

The maximum gains relative to umbo motion were 59.5 and 53.7 dB at the sites with CF of 5.5 and 14.5 kHz, respectively. Throughout this work, those numbers will be considered as the maximum gain of the cochlea. At the location with CF=5.5 kHz, the maximum gain relative to stapes was 69.8 dB (no stapes recordings were performed in the other cochlea). The sharpness of the cochlear filter was measured using the Q_{10} factor (tip frequency divided by its 10-dB bandwidth) and it was found to decrease as a function of click level. In the 14-kHz region, for example, Q_{10} was 3.22 at 54 dB SPL and 1.43 for 114 dB SPL clicks. Similar values of Q_{10} were estimated in the 5.5-kHz region (3.28 for 46-dB SPL clicks and 1.24 at 116-dB SPL).

The first column of Fig. 6 shows BM phases relative to umbo phases obtained from responses to click stimuli of various levels. Figure 6(a) displays response phases for the recordings at CF=14.5 kHz, whereas Fig. 6(b) shows results from the 5.5-kHz site. Square symbols mark the phase lag relative to umbo motion at CF. Note that in spite of the obvious changes as a function of level in gain function shown in Fig. 5, intensity-dependent phase changes are relatively small (see below).

Another way of displaying phase data is by plotting the group delay as a function of frequency (Allen, 1983). For a

linear system with a frequency response expressed as $H(\omega) = A(\omega)e^{j\varphi(\omega)}$, the group delay is defined as (Papoulis, 1977),

$$t_g = -\partial\varphi(\omega)/\partial\omega.$$

If the input envelope, $f(t)$, to a system varies slowly, the group delay represents the envelope delay at the output of the filter [i.e., $f(t-t_g)$]. Group delays were estimated from phase versus frequency curves of responses to clicks, such as those shown in Figs. 6(a) and (b). For frequencies around CF, group delays are dependent on the level of stimulation, as they generally decrease as a function of level. Note also that group delays versus frequency curves have maximum values at frequencies above CF. In the 14-kHz region [Fig. 6(c)], measured CF group delays were 0.361, 0.408, 0.3, 0.337, 0.383, 0.431 and 0.455 ms for levels 114, 104, 94, 84, 74, 64, 54 and 44 dB SPL, respectively. In the 5.5-kHz region [Fig. 6(d)], CF group delays were 0.963, 0.994, 0.955, 0.933, 0.953, 1.006 and 1.035 ms for levels 116, 106, 96, 86, 76, 66, 56 and 46 dB SPL, respectively. Group delays in the hook region tend to be smaller than those measured at the 5.5-kHz region by an average difference of 588 μs. The difference in delays cannot be explained by the cochlear delay differences (20 μs, Sec. III A) between the two sites or by their respective quality factors; such a difference probably reflects dispersion, or deviation from scaling, in the cochlea. At low enough frequencies ($<$ 0.5 CF), group delays appear to be frequency- and level-independent.

D. Longitudinal spatial patterns

Figure 7(a) shows BM responses to same-level clicks at seven locations in the same cochlea, with the bottom waveform displaying responses from the lowest CF (i.e., the most apical site). Onset delays measured from each response are 224, 224, 224, 228, 228, 232 and 232 μs, at positions whose CFs are 13.7, 12.7, 11.7, 10.5, 9.3, 7.1 and 6.5 kHz, respectively. The measured distance between the most basal and the most apical sites was 2.1 mm. Based on Greenwood's formula (Greenwood, 1990) for the chinchilla, the cochlear distance between the most basal (CF=13.7 kHz) and the most apical (6.1 kHz) sites is 2.79 mm. From the above measurement (=2.1 mm), the speed of the traveling wave in the base of the cochlea was estimated at 262.5 m/s, which is much less than the speed of sound in water but close to the speed of sound in air. A second set of recordings in another cochlea is shown in Fig. 7(b). The five locations span 1.2 mm with a CF range between 10.5 and 15.4 kHz. Using Greenwood's map, the distance between the two extreme points was estimated at 1.44 mm. Starting at the most basal location, onset delays were 224, 228, 232, 232, 232 μs. The speed of the traveling wave can be estimated at 150 m/s. The discrepancy between the two estimates of traveling wave velocity might originate from the time resolution of these recordings (4 μs). Table I shows onset delays measured in the middle ear and two cochlear locations in eight ''healthy'' (i.e., nonlinear) preparations. Cochlear delays (measured with respect to the umbo) at the 9-kHz locations were 33.6 $\pm$6.7 μs (mean $\pm$ standard deviation). Shorter cochlear delays were found in the 11–14.5-kHz region (24$\pm$2.8 μs).

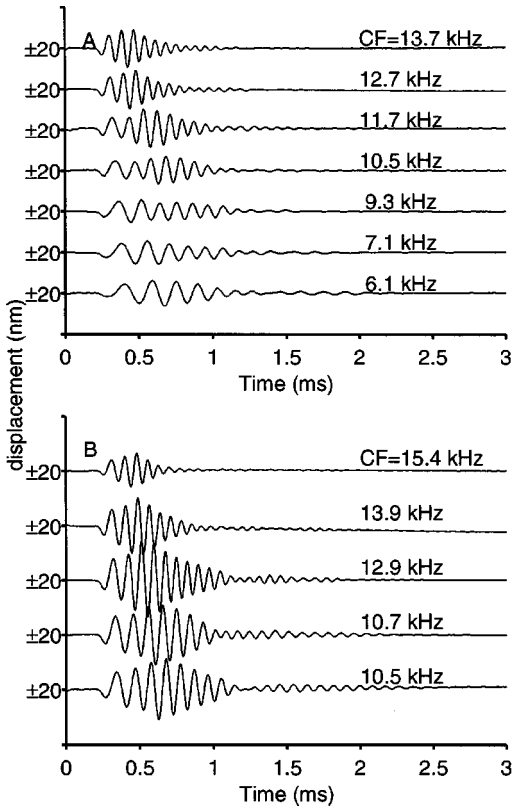


FIG. 7. Basilar membrane responses to a condensation click at several locations in the same preparations [CB31 in panel (a), CB30 in panel (b)]. Level of stimulation=87 dB SPL.

Amplitude and phase transfer functions for each of the recordings shown in Fig. 7 were also obtained. Figure 8(a) shows the amplitude of the 10-kHz component for several stimulation levels at each of the locations of Fig. 7(a). The spatial distribution of the 10-kHz component shows a shift with level toward the 0.3-mm location (CF=12.7 kHz). The maximum vibration evoked by the click at the 10-kHz component did not occur at the place in the basilar membrane with CF=10 kHz (indicated by an arrow whose position was determined using Greenwood's map), but at a location about 1 mm basal to it. Similar results were obtained from the results of Fig. 7(b) for a 12-kHz component. That is, the 12-kHz component reaches a maximum displacement during high-level stimulation at a location basal to the place of its CF.

TABLE I. Onset delays measured at the middle ear (ME) and at two cochlear locations in eight nonlinear preparations. ME onsets were measured at the umbo, with the exception of CB54 and CB58, in which onsets were measured at the stapes, near the incudostapedial joint.

Animal	ME onset (μs)	CF_1 (kHz)	BM onset (μs)	CF_2 (kHz)	BM onset (μs)
CB21	196	9	232	14.5	220
CB24	208	9	236	11	232
CB27	216	9	260	14	236
CB28	212	9	240	11	236
CB29	200	9	232	12	228
CB54	204	6.5	224	15	212
CB56	220	6.3	240	7.3	240
CB58	204	5.5	252	6.5	244

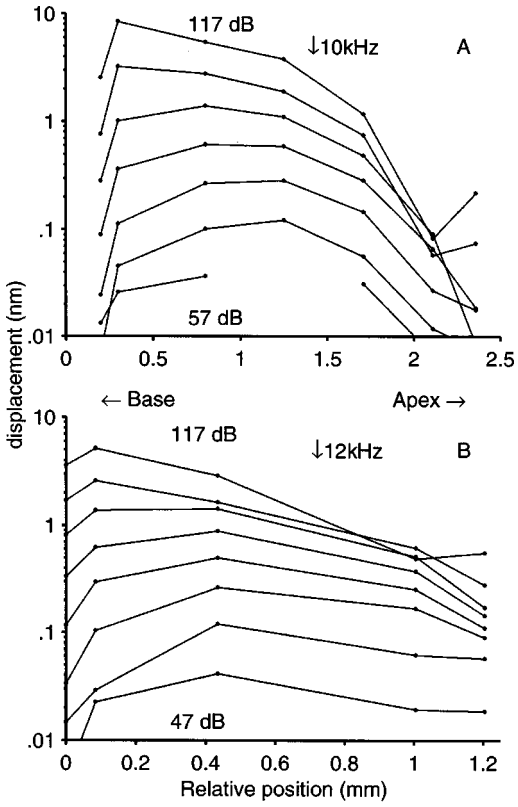


FIG. 8. BM displacements at 10 kHz (a) and 12 kHz (b) obtained from Fourier transforms of responses to clicks at seven (a) and five (b) positions along the basilar membrane at several levels (10-dB steps). Data in panel (a) from animal CB31 and those in panel (b) from animal CB30. Radial locations, measured from the spiral ligament end of the basilar membrane, for the beads in panel (a) are approximately (from base to apex): 80, 120, 140, 120, 100, 120 and 100 μm. For sites shown in (b), the locations are (measured from the spiral lamina): 180, 140, 130, 150 and 100 μm. Locations of the recording sites were measured as described in Sec. II. The positions of the arrows, however, were estimated using Greenwood's map.

Figure 9(a) shows the response phases relative to umbo motion, obtained from the data shown in Fig. 7(a). These relative BM phases show shallow segments at low frequencies, followed by steeper segments at frequencies around CF. Phase lags at CF (square symbols) appear to be independent of the CF of the recording site. From the phase functions shown in Fig. 9(a), we estimated group and phase delays. Group delays obtained from the results in Fig. 9(a) show a maximum value at frequencies just above CF [Fig. 9(b)]. As was the case for the results shown in Sec. III C, the differences between group delays measured in the hook and those measured in more apical regions cannot be accounted for by the difference in cochlear delays ($\leqslant 8$ μs, see above). For a linear system, such as the one described in Sec. III C, phase delay is defined as (Papoulis, 1977),

$$t_p = -\varphi(\omega)/\omega.$$

For a tone burst passing through a linear filter, the phase delay represents the delay of the carrier [i.e., $\cos\omega(t-t_p)$; Papoulis, 1977]. Phase delays [Fig. 9(c)] were evaluated at seven frequencies for each of the phase curves shown in Fig. 9(a). These results show that at the BM locations where recordings were obtained, the phase delay of the carrier is proportional to its frequency. The velocity, v, of a given frequency component, ω, of the traveling wave at a given location can also be estimated from phase delays, by obtaining the ratio of the distance between two adjacent locations (1 and 2), Δs, and the phase delay difference between responses at the two sites, Δt_p,

$$v(\omega) = \frac{\Delta s}{\Delta t_p} = \frac{\Delta s * \omega}{\varphi_1(\omega) - \varphi_2(\omega)}.$$

Plots of velocity versus position along the BM for four frequencies are shown in Fig. 9(d). In general, the velocity of

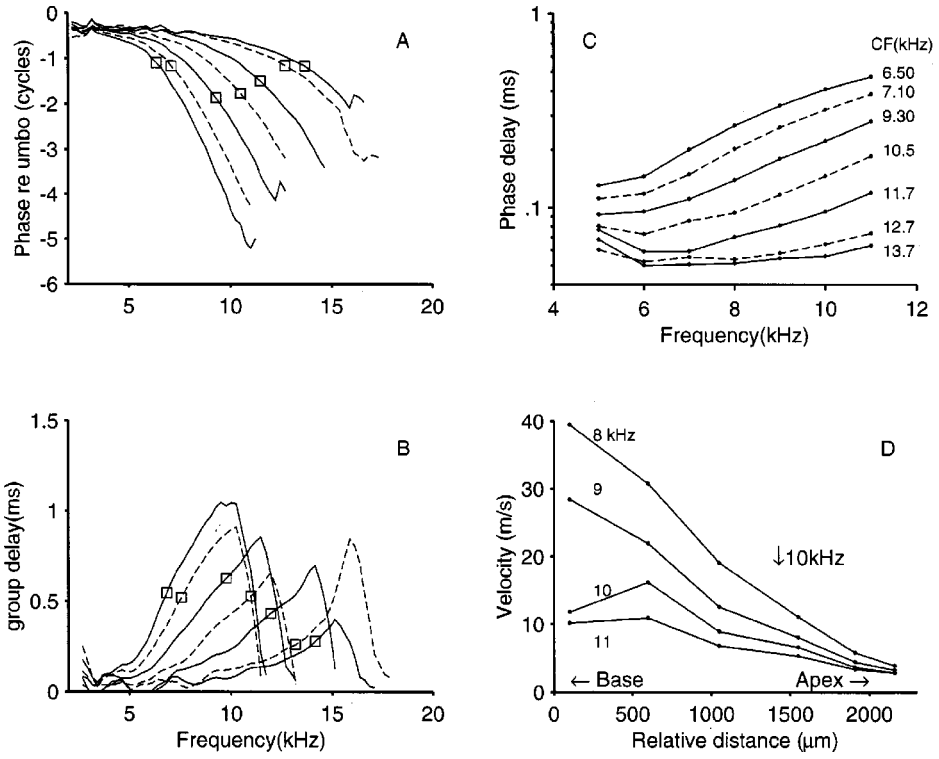


FIG. 9. (a) Basilar membrane phases (upper panel, phase of scala vestibuli re inward movements of umbo) obtained from responses shown in Fig. 8. (b) Group delays as a function of frequency obtained from data shown in (a). Rectangles indicate the phase or group delay at CF. (c) Phase delays versus frequency curves obtained from each of the curves shown in (a). (d) Velocity versus location curves obtained from phase delay data displayed in (c).

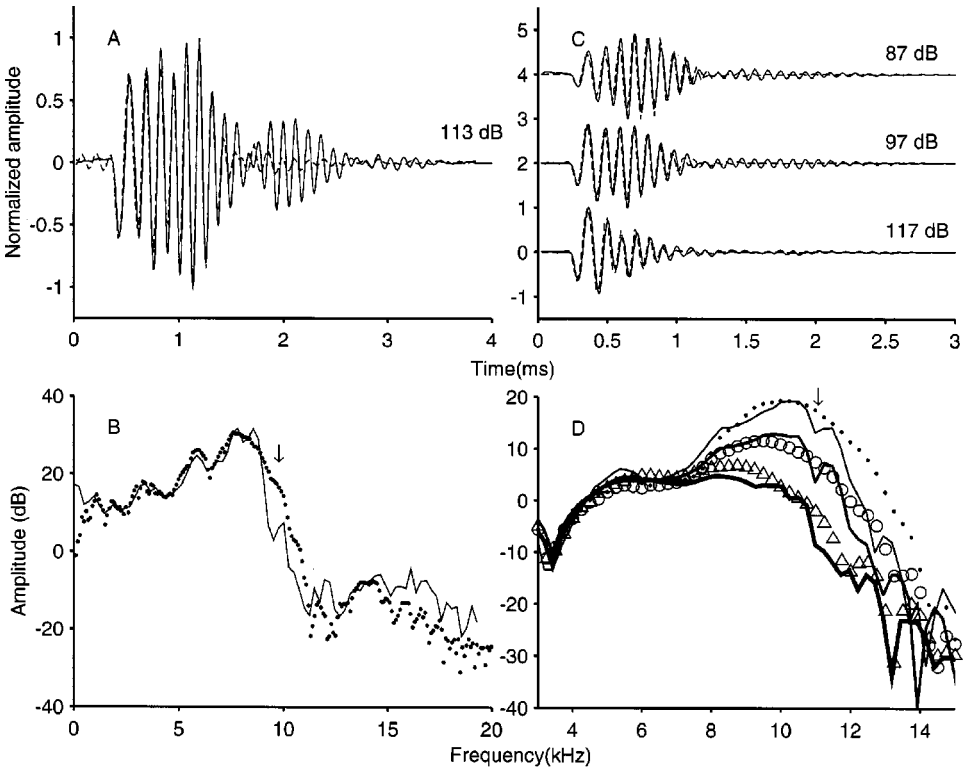


FIG. 10. [(a) Animal CT06] and [(c) animal CB30] Basilar membrane responses to clicks (continuous lines) and ''synthetic'' impulse responses obtained from BM responses to tones (dashed lines) in two cochleae. Panels (b) and (d) display Fourier transform amplitudes (continuous lines) of the click responses shown in (a) and (c), respectively. Also plotted using symbols [dots in (b); dots, circles and squares in (d)] are BM responses to tones.

the wave decreases as it travels down the cochlea even before reaching its point of resonance.

E. Comparisons of BM responses to clicks and tones

BM responses to clicks and tones were compared in both the time and frequency domains. Comparison in the time domain involved the creation of a ''synthetic'' impulse response, obtained by taking the inverse Fourier transform of the steady-state tone response data. As mentioned in the methods section, tone and click stimuli were presented without being compensated by the acoustic calibration. Figures 10(a) and (c) show measured click responses (continuous lines) and synthetic impulse responses (dashed lines) in two cochleae. All measured responses show the ''two-lobe'' wave shape, which is characteristic of many nonlinear cochleae. The ''second'' lobe, or ringing epoch, occurring after 1–1.5 ms, seen in the measured click responses (continuous lines), however, does not appear in the synthetic waveforms (dashed lines). Throughout the duration of the first lobe of oscillation (in results not shown), the instantaneous frequencies of click and synthetic impulse responses are very similar.

Figure 10(b) shows the Fourier transform amplitude (continuous line) of the click response shown in panel A and the tonal response spectra (dots). Figure 10(d) shows similar results from another cochlea for clicks and tones at three stimulus levels (line thickness is in proportion to click level). At the lowest stimulus level [Fig. 10(d)], amplitude responses for both clicks and tones show maximum values at a frequency around 10 kHz. As the stimulus level increases, the position of the amplitude peak shifts to lower frequencies. Although there is an overall good match between click responses transforms and tonal response spectra, particularly at the highest level of stimulation (triangles in panel D), the responses differ in two aspects. Near-CF notches, which appear in click responses at around 9 kHz [Fig. 10(b), arrow] or 11 kHz [Fig. 10(d), arrow], are not shown in the responses to tones. High-CF notches, such as the one around 13 kHz in Fig. 10(d), appear in both tones and click responses. Also, the amplitudes of the click-response transforms in a frequency range above CF are generally smaller than those obtained in responses to tones. For the results in Fig. 10, it is apparent that the match between the spectra of click-response transform and tonal responses is better for frequencies below the frequency corresponding to maximum displacement than for frequencies above this point.

IV. RESPONSES TO SCHROEDER-PHASE COMPLEX STIMULI

BM responses to another type of broadband stimuli, the Schroeder-phase complex, were also obtained. Using a linear BM model developed by Strube (1985), Smith *et al.* (1986) showed that depending on the starting phase of the complex, BM responses to the Schroeder-phase complex showed differences that could explain the results of their psychophysical experiments. As discussed in Sec. V, some BM properties that are easily visualized in responses to clicks, such as the ''glide'' in the response (Fig. 4) or the shape of the phase versus frequency response curve (Fig. 9), can probably explain the differences between BM responses to the +SCHR and −SCHR stimuli.

A. Response waveforms and Fourier analysis

Figure 11 shows BM responses to the +SCHR stimulus (panel A) and the −SCHR stimulus (panel B) at two levels in a preparation with CF=8 kHz. At both levels, responses to

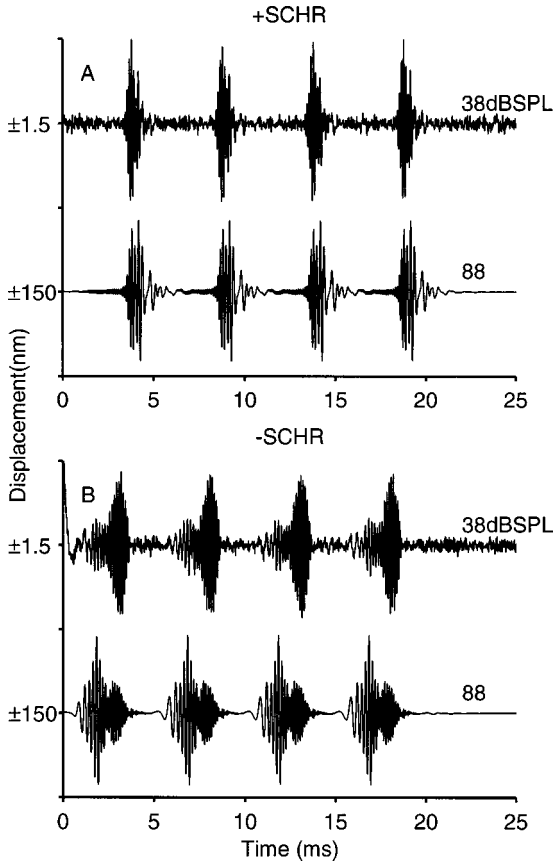


FIG. 11. BM responses to +SCHR (a) and −SCHR (b) stimuli at two levels. Animal=CT06, CF=8 kHz. Numbers along *y*-axes indicate displacement between adjacent tick marks.

the +SCHR stimuli contain valleys, or dips, that are wider than responses to −SCHR stimuli. As a consequence of this, responses to +SCHR stimuli are ''peakier'' than those evoked by the −SCHR stimulus, i.e., responses to +SCHR stimuli are more compact than responses to −SCHR stimuli. Responses evoked by −SCHR stimuli contain larger amounts of spectral energy than those measured in responses to +SCHR sounds. For example, the rms (root mean squared) value of the responses to the 88-dB −SCHR stimulus is 34 nm, whereas responses to the +SCHR stimulus at the same level have an rms=30 nm. Because of the cochlear nonlinearity and changes in the stimulus instantaneous frequency, the peak value in responses to −SCHR stimuli occurs at different times, depending on the level of stimulation. The lower the stimulus level, the later the peak value occurs, indicating a change with level in the frequency selectivity of the cochlea. For example, whereas responses to the 88-dB −SCHR stimuli reach their peak values when the instantaneous frequency of the response is about 5.05 kHz, responses to the 38-dB stimuli reach their peaks when their instantaneous frequency is around 8.6 kHz. On the other hand, in responses to the +SCHR stimulus, maximum values always occur at similar times, yet the instantaneous frequencies at the responses peaks are similar to those obtained in responses to −SCHR stimuli (5.1 kHz for the 38-dB stimuli and 8.08 kHz at 88 dB). It is also possible to see the effect of BM compression by comparing the growth in amplitude in the responses, roughly 40 dB, to the difference in stimulation level, 50 dB.

Figures 12(a)–(d) show the responses to the same stimuli in two other cochleae. Each response waveform was normalized to its maximum value. The same characteristics found in the responses shown in Fig. 12 are present in the responses of these other cochleae.

Figures 13(a)–(d) display the Fourier transform amplitude of BM responses to −SCHR (left columns) and +SCHR stimulus (right columns) at two intensity levels. BM responses are those shown in Fig. 11. Superimposed on Figs. 13(c) and (d) dashed lines, are the envelopes of the Fourier amplitudes from responses to −SCHR stimulus [Figs. 13(a) and (b)]. Whereas time domain responses to +SCHR and −SCHR stimuli of the same intensity look very different (Figs. 11 and 12), in the frequency domain, the responses are relatively similar. For the 88-dB SPL stimulus, the response maximum occurred at approximately 5.5 kHz (regardless of the stimulus used, i.e., +SCHR or −SCHR), and shifted toward higher frequencies as the levels of stimulation decreased (around 8 kHz for the 38-dB SPL stimulus).

Figures 14(a)–(d) show the responses to the Schroeder-phase complex stimuli in two damaged preparations. Data in the left column come from a linear site around the 8-10-kHz region. Because of the linear behavior, BM responses to different stimuli levels are scaled versions of each other. It is not clear whether responses to the +SCHR stimuli are ''peakier'' than those obtained in response to the −SCHR stimuli (as in the case of a nonlinear preparation). Data in the right column were recorded in the hook region and display some degree of nonlinearity, as evident from changes in wave shapes of responses as a function of level. Responses to +SCHR stimuli appear to last longer in linear and damaged preparations [Figs. 14(a) and (c)] than in the responses measured in healthier cochleae.

B. Analytic signal representation of responses to the Schroeder-phase complex

Because of the time-varying characteristics of BM responses to the Schroeder-phase complex, traditional Fourier analysis, such as the one done in Sec. IV A, does not reflect the dynamical aspect of such responses. We therefore computed both the envelope, to estimate the ''duration'' of the responses, and the instantaneous frequency of BM responses to such stimuli. Figures 15(a) and (b) show the envelope of the responses to −SCHR (continuous lines) and +SCHR (dashed lines) at two levels [88 and 58 dB, in panels (a) and (b), respectively]. Envelopes were truncated for values below 20% of their maximum, since the envelopes were too noisy in this range. Envelope duration was defined as the delay between the first time the envelope exceeded 20% of its maximum value and the last time this occurred, within one period of the stimulus waveform. This analysis reveals that the ''width,'' or duration, of the envelope is larger for responses to −SCHR than for responses to +SCHR stimuli. Figure 15(c) shows envelope duration curves for responses to +SCHR stimuli (filled symbols) and −SCHR stimuli (open symbols) as a function of stimulus level obtained from

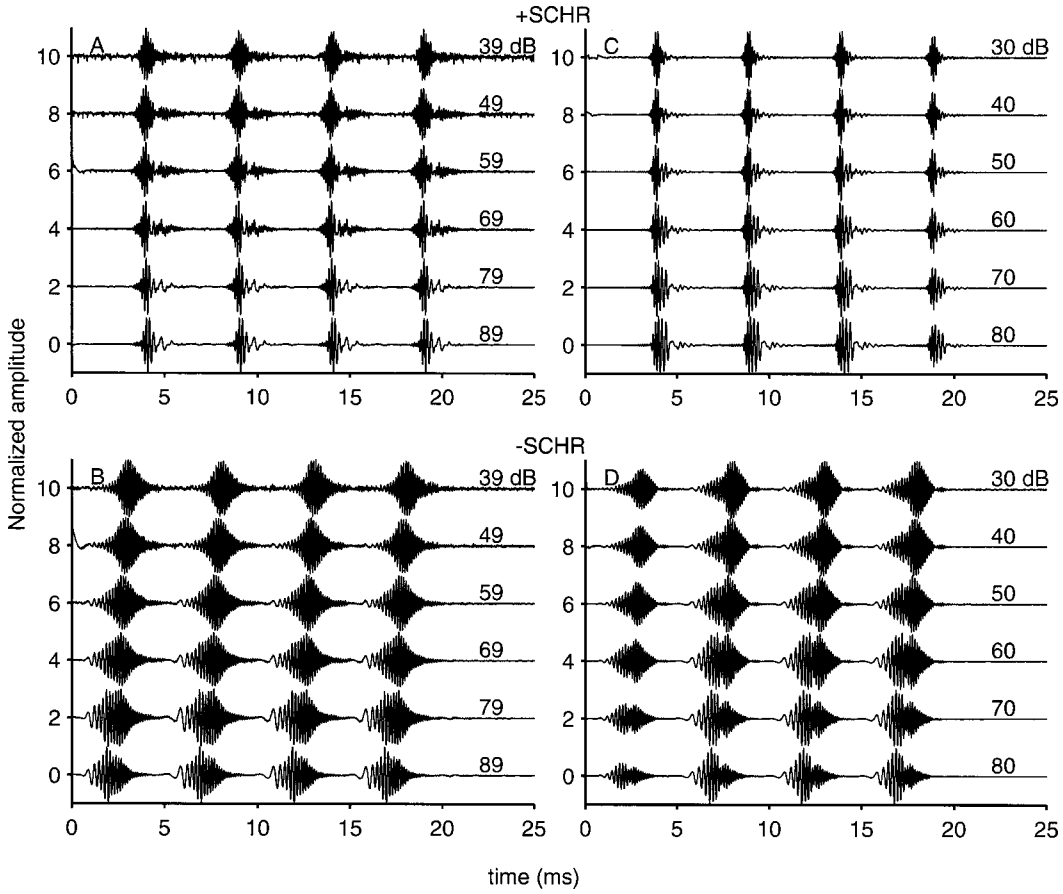


FIG. 12. BM responses to +SCHR (a) and (c) and −SCHR (b) and (d) stimuli at several levels in two cochleae. Animals CT17 [(a) and (c), CF=8 kHz] and CB61 [(b) and (d), CF=7.5 kHz]. Amplitudes were normalized to the response's peak.

recordings in three cochleae. Consistent with the impressions obtained from the upper two panels, there is a significant difference in the duration of both sets of responses. Up to certain levels, these results also show slight increases in the envelope duration with increasing stimulus level.

Instantaneous frequencies were obtained from one period of the responses to +SCHR stimuli [Figs. 16(a)–(c), thick dashed lines] and −SCHR stimuli [Figs. 16(a)–(c), thick continuous lines]. Glitches in the instantaneous frequency curves of the responses to the +SCHR stimulus (at around 4 ms) are a consequence of small values, or dips, in the response waveform. Stimulus level was 88, 73 and 58 dB SPL for the results in Figs. 16(a), (b), and (c), respectively. The instantaneous frequencies of these two stimuli are also shown in the same figure (thin lines). Note that the speed of the stimulus sweep is rather high, 3.845 MHz/s. The slope, or speed, of the instantaneous frequency versus time curves obtained from responses to the −SCHR stimulus appears lower than those of the stimuli. This difference indicates that the BM can follow the instantaneous frequency of the stimulus when the frequency decreases (+SCHR) better than when the instantaneous frequency increases (−SCHR). In other words, it takes more time for the BM to follow the instantaneous frequency from 5 kHz to a value equal to CF (=8 kHz), than when the instantaneous frequency of the stimulus goes from CF to 5 kHz.

Figures 17(a)–(c) display the envelopes, instantaneous frequencies (plotted as a function of time) and envelope duration (as a function of level) for the data from the damaged cochlea in Figs. 14(a) and (b). Envelope durations for responses to +SCHR (dashed lines) and −SCHR stimuli (continuous lines) are longer than those measured in nonlinear preparations [Fig. 15(c)] and remain constant with changes in the stimulus level [Fig. 17(c)]. Results from a slightly nonlinear preparation [Figs. 17(d)–(f)] show envelope durations that are longer than those measured in the more nonlinear preparations but do not remain constant as a function of level. Visual examination of instantaneous frequency versus time curves obtained in damaged preparations [Figs. 17(b) and (e)] reveals that the absolute values of the slopes of those curves (obtained in responses to −SCHR and +SCHR stimuli) are similar, more so than in the case of nonlinear preparations (Fig. 16). In other words, instantaneous frequency curves of responses to −SCHR and +SCHR stimuli in linear preparations [Figs. 17(b) and (e)] appear to follow the instantaneous frequency of the stimulus at similar speeds.

Instantaneous frequency of BM responses to +SCHR and −SCHR stimuli were further analyzed by estimating the amount of time it takes them to go between 5 and 10 kHz in four preparations (Fig. 18). The delay between the time the response first reaches 5 kHz and the time to reach 10 kHz in response to −SCHR stimuli [open symbols, Fig. 18(a)] is usually larger than the delay measured in responses to +SCHR stimuli [filled symbols, Fig. 18(a)]. Figure 18(b)

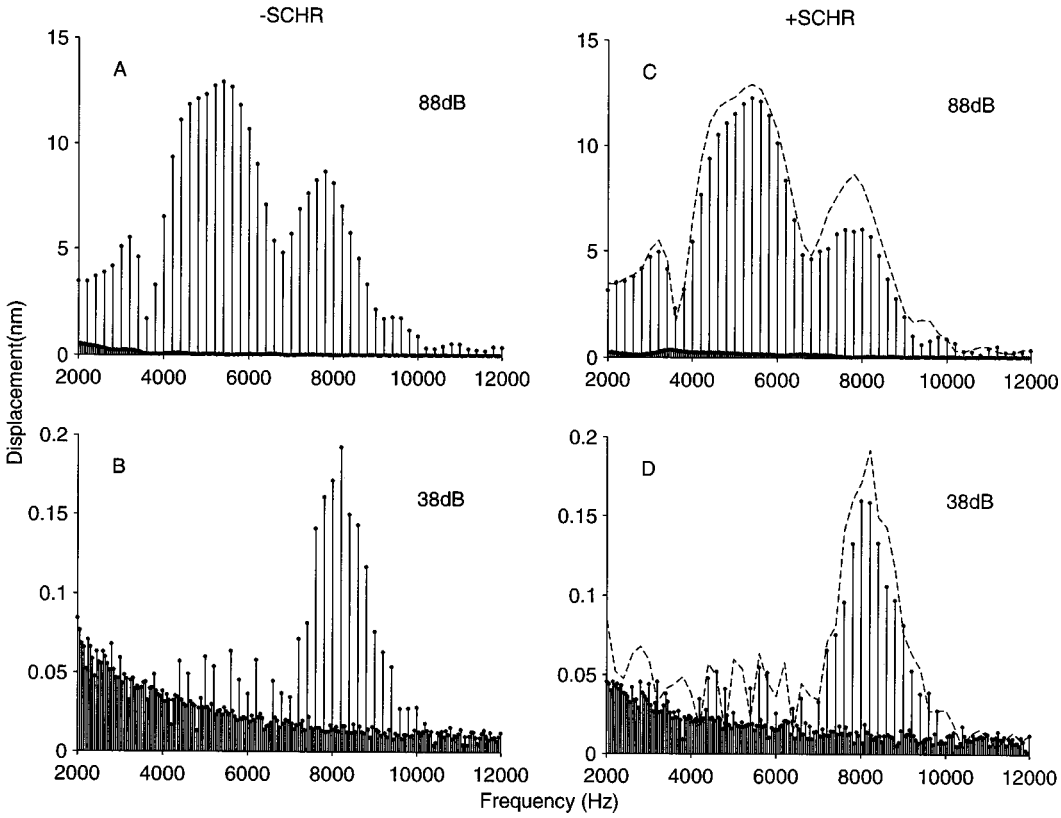


FIG. 13. Fourier transform amplitudes of responses to +SCHR (c) and (d) and −SCHR stimuli (a) and (b) at two levels. Fourier analysis was performed using DFTs (discrete Fourier transforms) at multiples of 50 Hz. Dashed lines in panels (c) and (d) represent Fourier amplitudes from panels (a) and (b) (at multiples of 200 Hz). Animal=CT06.

shows the difference between the times indicated by open and filled circles. For healthier preparations [triangles, squares and circles in Fig. 18(b)], the difference between those delays decreases as a function of level. That is, the nonlinear processing in the basilar membrane makes responses to +SCHR and −SCHR stimuli differ more at lower levels, when gain of the cochlear amplifier is the greatest.

C. Suppression of CF tones by Schroeder-phase complex stimuli

Suppression effects of Schroeder-phase complexes were studied by simultaneously playing a tone and a +SCHR or a −SCHR stimulus. The amount of suppression was defined as the reduction in the amplitude of the tone in the presence of the Schroeder-phase complex throughout the entire response duration. Figure 19(a) shows one cycle (5 ms) of the BM responses to a +SCHR stimulus (92-dB SPL) and a near-CF tone (8.1 kHz) at two levels (76-dB, continuous lines, and 6-dB SPL, dashed lines). The frequency of the tone (8.1 kHz) is not present in the harmonic complex. The two responses are virtually identical in the later part of the responses (between 8.5 and 10 ms), i.e., where the response to the +SCHR alone reaches its peak value. The response to the 76-dB tone is only obvious during the response ''valley.'' The same can be said about the responses to −SCHR stimuli and a tone [Fig. 19(b)]. We also subtracted the response to suppressor + tone (6 dB) from the responses to suppressor + tone (76 dB) and the results are shown in Fig. 19(c). Because responses to the 6-dB tone are very small, waveforms in Fig. 19(c) represent an approximation to the BM response to the 76-dB tone (in the presence of a suppressor). [This was done because we did not measure responses to the SCHR alone when two microphones were used.] Suppression in the response is more prominent during response peaks, when the basilar membrane is responding mostly to the suppressor, i.e., Schroeder-phase complexes evoked a temporal, or phasic, suppression. Figure 19(c) also shows that the amount of suppression by the −SCHR stimulus is larger than the one obtained using the +SCHR stimulus, as determined by the width of the ''gaps'' in the responses and the rms values of the waveforms (9.11 and 6.82 nm for responses suppressed by +SCHR and −SCHR stimuli, respectively).

The response of the basilar membrane to an 8.1-kHz tone was measured alone (Fig. 20, thin solid lines with filled symbols) and in the presence of −SCHR stimuli [Fig. 20(a)] and +SCHR stimuli [Fig. 20(b)] at several levels. The thickness of the lines is proportional to the level of the suppressor. The amount of suppression increased as a function of the suppressor level, with the highest effects occurring at the lowest levels of the probe signal. It can also be concluded that the −SCHR stimulus is more effective as a suppressor than the +SCHR stimulus. Figure 20(c) shows the response of the BM to a 36-dB SPL 8.1-kHz tone presented simultaneously with a Schroeder complex relative to the BM response to the tone alone. When the −SCHR stimulus was also present, the response to the 8.1-kHz tone decreased by as much as 23 dB, which is approximately 10 dB more than

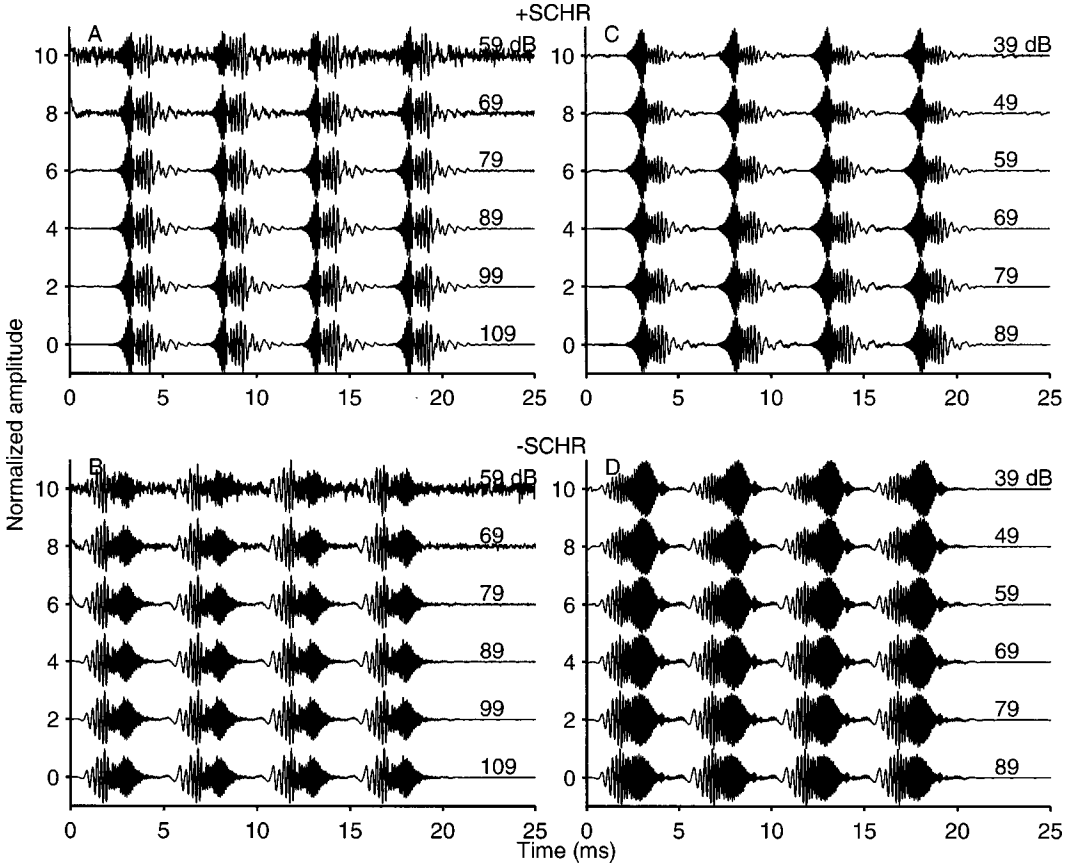


FIG. 14. BM responses to Schroeder-phase stimuli in a linear (a) and (b) and a slightly nonlinear preparation (c) and (d). Panels (c) and (d) come from responses recorded from a site in the hook region. Animals CB64 (a) and (b) and CB63 (c) and (d). Recordings were performed without a coverslip.

in the presence of the +SCHR stimulus. Similar results were obtained in another cochlea. In summary, the −SCHR stimuli produced larger amounts of suppression than did the +SCHR stimuli.

V. DISCUSSION

A. Responses to clicks

The data reported in this paper indicate that BM click responses in the hook region behave similarly to responses in a more apical region (5–9 kHz). In our best preparations, the maximum BM gain relative umbo motion was 55–60 dB at the sites corresponding to CF=5.5 and 14 kHz. Relative to stapes, this gain increases by about 10 dB. Sharpness of the cochlear filter, measured using the quality factor, Q_{10}, also is quite similar, in our best preparations, in both hook and 5–9-kHz regions. All the click responses in the basal region under study show frequency modulation, easily observed in instantaneous frequency versus time curves. The amount of frequency modulation (as determined by the time constant of instantaneous frequency versus time curves) is not constant along the base of the cochlea. It was found that the higher the CF, the greater was the amount of modulation (slower time constants).

BM responses to clicks show an asymmetry in the frequency domain in both amplitudes and phase components. Asymmetry around the peak of the magnitude response can be seen in Fig. 5. Group delays shown in Figs. 6 and 9 display a maximum value at a frequency that does not correspond to the frequency of the largest peak of the amplitude response, a consequence of phase asymmetry. When the peak of the group delay occurs at the same frequency as the maximum amplitude response, such as in the case of gammatone filters, frequencies below and above CF arrive together and average out (Lyons, 1997), with only the period corresponding to CF being present in the click response. In general, the impulse response of an asymmetrical filter is an amplitude- and phase-modulated signal (Papoulis, 1977). In the case of the cochlea, it is thought (Lyons, 1997) that the asymmetry in phase delay versus frequency curves translates in the time domain as a ''glide,'' i.e., an impulse response with an instantaneous frequency varying from below CF to approximately CF.

B. Cochlear delays

Three types of delays were estimated from BM responses to clicks: cochlear, group, and phase delays. The mathematical definition of each of these delays comes from linear system theory, for which responses to clicks and tones are related to each other via Fourier analysis. For nonlinear systems like the cochlea, delays obtained using clicks' responses are not necessarily identical to those obtained using tones. Results from Figs. 6 and 10, however, indicate that responses to tones and clicks are relatively close to each other. Delay estimates obtained using clicks, therefore, should be similar to those found using tones.

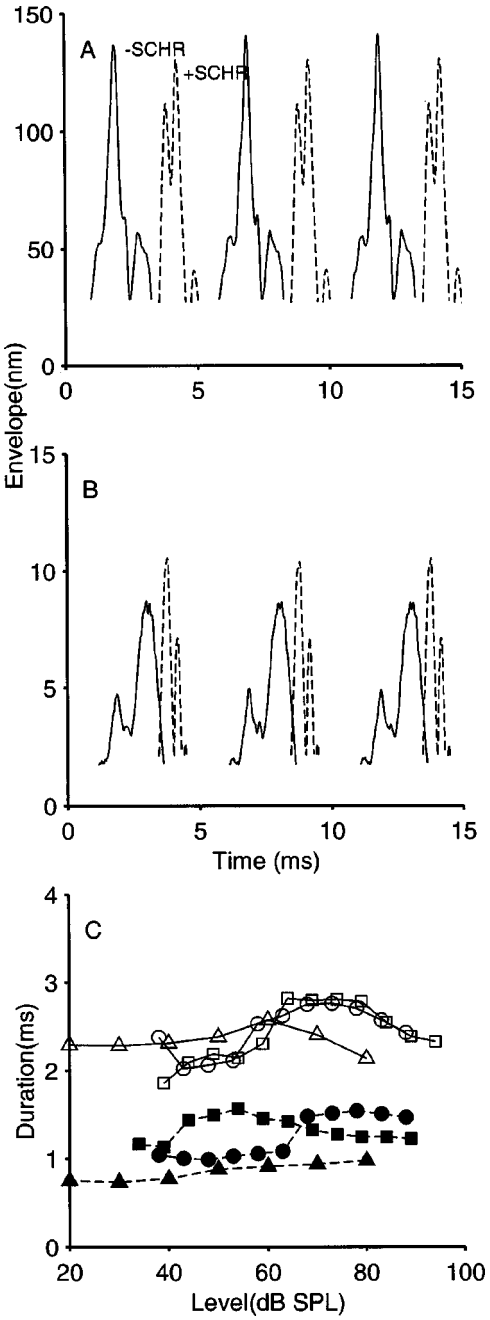


FIG. 15. Envelope (a) and (b) and envelope duration versus level curves (c) obtained from responses to −SCHR [continuous lines, open symbols in (c)] and +SCHR stimuli [dashed lines, closed symbols in (c)]. Data in panels (a) (level=88 dB) and (b) (level=58 dB) come from animal CT06. Data in panel (c) come from three animals (CT06, CT17, and CB61).

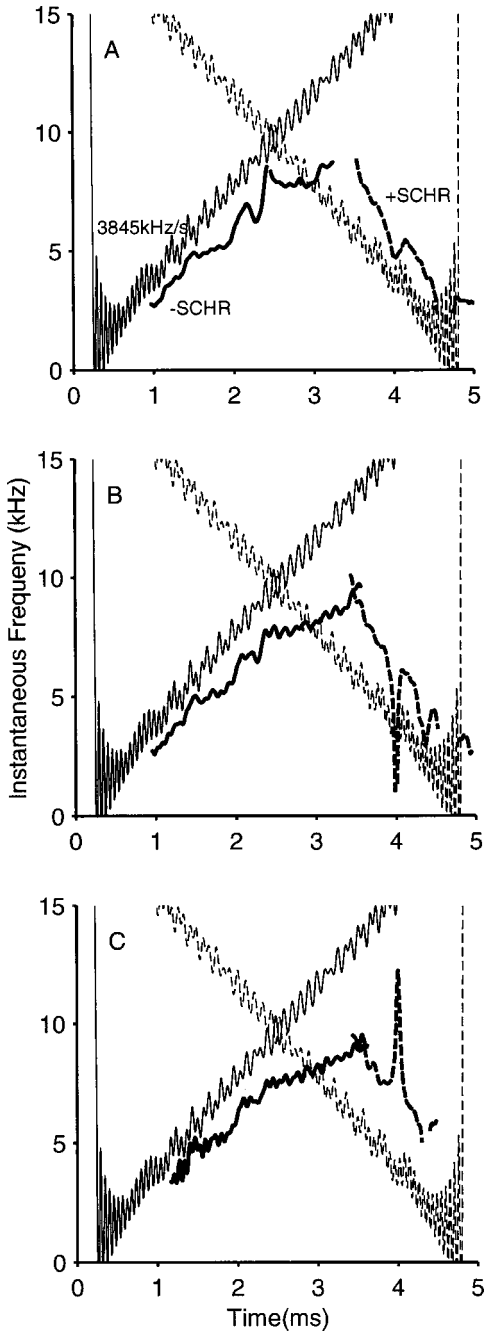


FIG. 16. Instantaneous frequency versus time curves for BM responses to −SCHR (thick continuous line) and +SCHR (thick dashed lines). Thin continuous and thin dashed lines display the instantaneous frequency of the stimuli. Stimulus levels=88, 73 and 58 dB for panels (a), (b) and (c), respectively. Animal=CT06, CF=8 kHz.

The definition of cochlear delay used in this work is equivalent to the one Békésy proposed for travel time (Békésy, 1949) and represents a pure (i.e., frequency independent) cochlear delay. The travel time can be obtained from the high-frequency slope of the phase versus frequency curve computed from BM click responses (Ruggero, 1980). We did not obtain travel times, or cochlear delays, from the phase versus frequency curves because of the lack of reliable phase information at high frequencies, hence we estimated them directly from the response waveforms. From travel time estimates the speed of the traveling wave was calculated. The two estimates (262.5 and 150 m/s) for the traveling wave velocity in the base of the cochlea are below the maximum speed limits predicted by Békésy's theory but higher than those expected from other theories, such as Zwislocki's (Wever and Lawrence, 1971). Velocities shown in Fig. 9(d) were obtained from phase delay estimates at a given frequency, ω_0, and represent the velocity of a frequency component of the click response along the cochlea. Hence, velocity estimates in Fig. 9(d) are frequency-dependent and are slower than the frequency-independent traveling wave velocity (i.e., the velocity estimates obtained from click onsets at two positions along the cochlea). In

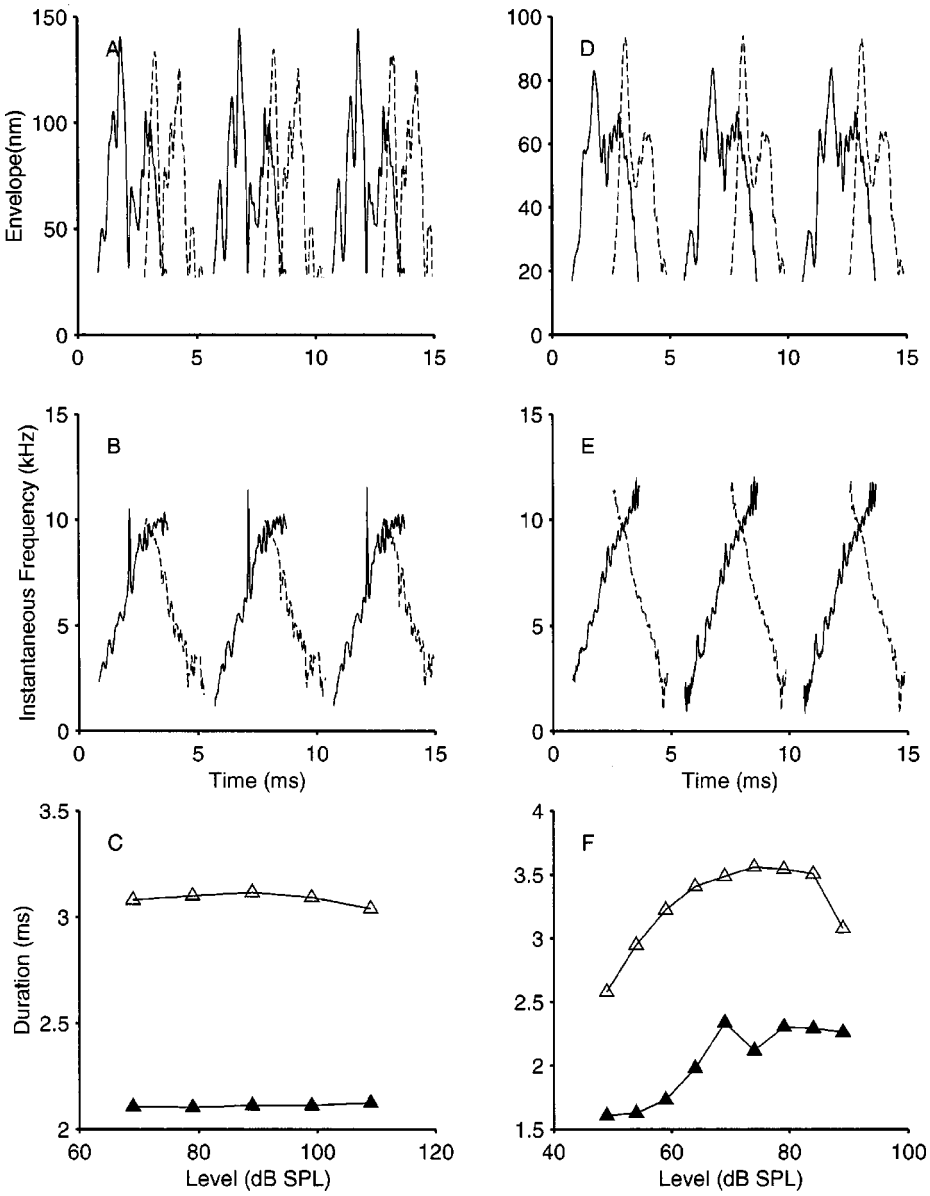


FIG. 17. Response envelope (a) and (d), instantaneous frequency versus time curves (b) and (e) and envelope duration versus level curves (c) and (d) obtained from BM responses to +SCHR (dashed lines) and −SCHR (continuous lines) stimuli. Data in the left column [panels (a)–(c)] come from CB64 and those in the right column (d)–(f) from animal CB63.

theory, the results in Fig. 9(d) and traveling wave velocities can all be generated from phase versus frequency curves of BM responses to clicks, the former by finding phase delays, $-\varphi(\omega_0)/\omega_0$, at several locations, the latter by computing $-\lim_{\omega\to\infty}\varphi(\omega)/\omega$, the travel time (Ruggero, 1980).

C. Spatial vibration patterns

Responses to clicks were measured at several locations along the basal region of the basilar membrane. Up to eight measurements with a CF range of more than one octave were obtained in a single cochlea. Plots of the spatial vibration pattern of a 10- or 12-kHz component show a shift with level toward a cochlear location basal to the CF site. This result agrees with amplitude versus frequency plots measured at a single point, which show a shift toward lower frequencies with stimulus level in the location at which the largest response occurs. It also agrees with a similar analysis performed using tones (Rhode and Recio, 2000). Our results, however, disagree with those obtained by Russell and Nilsen (1997), in which it was concluded that the location of maximum displacement for a given frequency does not change with level. The disagreement between the two studies is difficult to explain, but we certainly do not believe that species differences (guinea pig versus chinchilla) play an important role. A possible explanation comes from the fact that the results of Russell and Nilsen [1997, Fig. 1(d)] originate not from one single cochlea but from many. Although Russell and Nilsen's recordings from around the CF region (=15 kHz) were derived from one cochlea, recordings from the most basal region (17–27 kHz) were obtained in separate preparations. The inclusion of those extra points is, in our view, the main source of the discrepancy. For example, at the 22-kHz site, in which displacements evoked by a 15-kHz tone are in the tail of the tuning curve, BM motion grows nonlinearly as a function of stimulus level. In fact, the overall compression (slope) of the displacement versus level curve at 22 kHz is ≈0.57 [Fig. 1(d), Russell and Nilsen]. Displacements in the tail region of BM tuning curves should grow linearly.

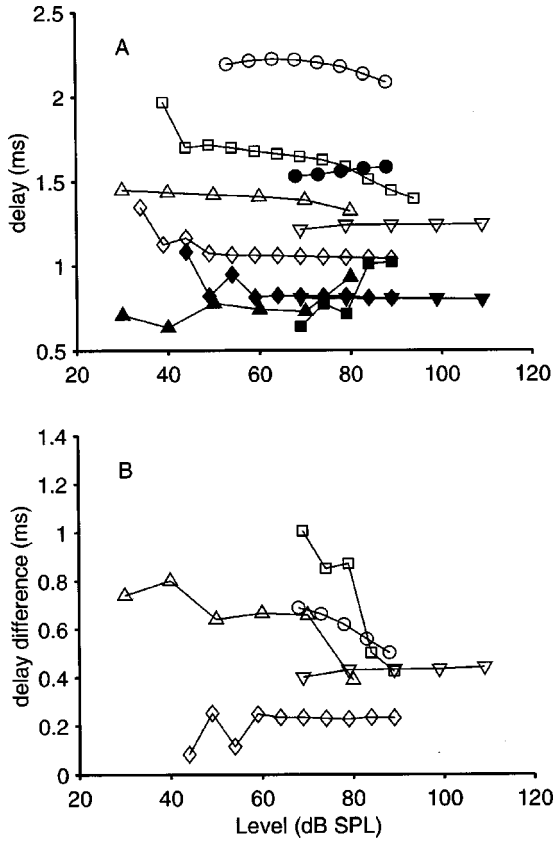


FIG. 18. (a) Time for an instantaneous frequency versus time curve to go from 5 to 10 kHz (open symbols, in response to −SCHR stimulus) and from 10 to 5 kHz (filled symbols, in response to +SCHR stimulus). (b) Time difference between open-symbol and filled-symbol curves. Each symbol represents data from a different preparation (○ CT06, □ CT17, △ CB61, ▽ CB63, and ◇ CB64).

D. Comparison of BM responses to clicks and tones

Previous comparisons of BM responses to clicks and tones (Robles *et al.*, 1976; Recio *et al.*, 1998) concluded that both responses match each other fairly well. Robles *et al.* (1976) noted that ''synthetic'' impulse responses, obtained by inverse Fourier transformation of steady-state sinusoidal responses, are very similar to BM click responses measured in the squirrel monkey. Recio *et al.* (1998) compared both responses in the frequency domain (i.e., by comparing the Fourier transform of BM click responses to the transfer function obtained from responses to tones) and concluded that responses to tones and clicks generally agree with each other. In that work, however, responses to tones were evaluated with 1000-Hz resolution, hence responses in the region just above CF, in which a notch is usually observed in responses to clicks (Recio *et al.*, 1998), were not obtained.

In the present study, we compared both responses in the time and frequency domain. In the time domain, synthetic click responses did not show the ''two-lobe'' waveform that usually can be observed, at certain levels, in BM responses to clicks (e.g., Fig. 10). Spectral notches observed just above CF (Recio *et al.*, 1998) were shown to be a consequence of a phase reversal between early and late frequency components associated with the second time domain lobe. In the frequency domain comparison, BM click responses showed a spectral notch in a region near CF, whereas BM tone responses did not.

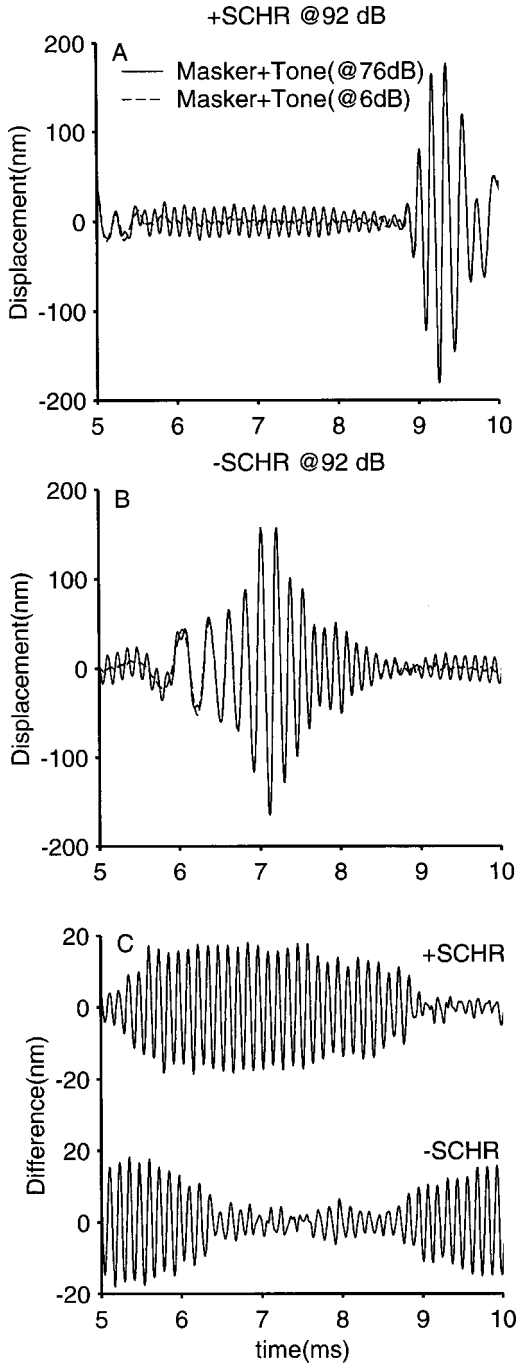


FIG. 19. (a) BM responses to the +SCHR and an 8.1-kHz tone at 72 dB SPL (continuous lines) and at 6 dB SPL (dashed lines). (b) Same as panel (a) but using the −SCHR stimulus as masker. (c) Responses to tone (at 6 dB SPL) +masker were subtracted from the responses to tone (72 dB) +masker. The masker type is indicated next to each trace. Animal CT17, CF=8 kHz.

Given the nonlinearity of BM motion, the existence of some differences between responses to tones and clicks was not unexpected. The fact that the ''two-lobe'' wave shape cannot be obtained from steady-state sinusoidal stimulation might be a consequence of the cochlear nonlinearity or perhaps of time-varying properties in BM motion.

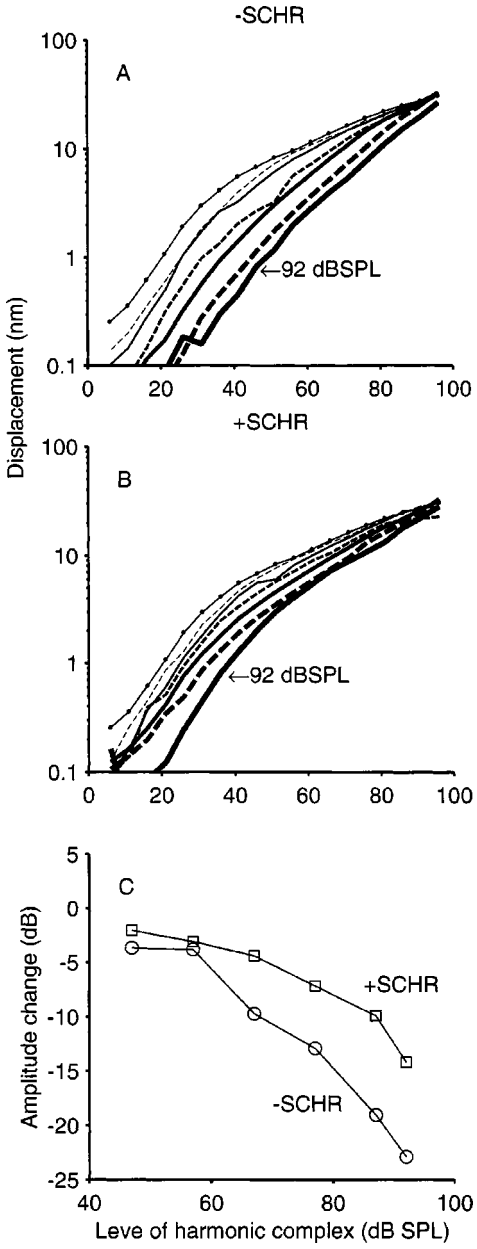


FIG. 20. Displacement amplitudes of BM responses to a near-CF tone [continuous lines with symbols, panels (a), (b) and (c)]. Dashed and continuous lines (without symbols) indicate displacement at the frequency of the near-CF tone in the presence of a −SCHR masker [panel (a)] and a +SCHR masker [panel (b)] of different levels (thickness of lines in proportion to intensity). Panel (c) shows a curve representing changes in the response to an 8.1-kHz tone when the harmonic complex is also present. Response amplitudes were obtained from the DFTs of the overall response. The level of the complex was varied between 47 and 92 dB SPL in steps of 10 dB (except between 87 and 92 dB). Square symbols indicate results using the −SCHR stimulus; circles show those obtained using the +SCHR stimulus. Animal CT17, CF= 8 kHz.

E. Responses to Schroeder-phase complex stimuli

BM responses to Schroeder-phase complexes were analyzed using the analytic signal representation, which allowed the estimation of the envelope and the instantaneous frequency of the responses. BM responses to +SCHR complexes are ''peakier'' than the waveforms obtained in responses to −SCHR stimuli. To a first approximation, responses to +SCHR appear to last less, for each of the stimulus periods, than responses to −SCHR stimuli. The envelope duration, which was used to quantify the duration of the response, confirms the previous assertion. This verifies simulations analysis done by several groups (e.g., Smith *et al.*, 1986), in which the stimuli were passed through a linear model of the basilar membrane. Envelope duration, however, does not remain constant across stimulus levels and appears to increase with level, reaching its highest values in linear preparations.

The behavior of BM responses to (+/−) Schroeder complexes is consistent with the theoretical analysis performed by Kohlrausch and Sander (1995), in which the curvature (i.e., the second derivative) of the phase versus frequency curve of the BM was used to account for the differences in the response to each of the two Schroeder phases. Curvatures of the phase versus frequency curves obtained from BM responses to clicks, such as those shown in Fig. 9(a), have a constant negative sign around most of the frequency response. The absolute value of the curvature is inversely related to CF. Whereas curvatures of −SCHR stimuli are of similar value and sign as those measured in BM responses to clicks, curvatures of +SCHR stimuli are of the opposite sign (but similar absolute value) to those measured in −SCHR stimuli. Hence, phase characteristics of BM responses to +SCHR stimuli show reduced curvatures (relative to the stimulus), and, as a result, peaked waveforms (Kohlrausch and Sanders, 1995).

The instantaneous frequency measured in responses to +SCHR stimuli differs from that obtained in responses to −SCHR stimuli, even in linear preparations. Measures of the instantaneous frequency in responses to the +SCHR stimulus reveal that those responses represent more accurately the rate of the instantaneous frequency of the stimulus. Responses to −SCHR stimuli have rates that are lower than those measured in the stimulus. Differences in instantaneous frequency slope in the responses also decrease with level.

When the Schroeder-phase stimulus was used as a suppressor, or masker, of a near-CF tone, the −SCHR stimulus produced greater suppression than did the +SCHR stimulus. It was shown (Fig. 19) that the largest amount of suppression occurs during the time in which the basilar membrane is responding to near-CF masker components. This and the fact that responses to −SCHR stimuli appear to last longer than those to +SCHR stimuli are a possible explanation of the better role of −SCHR stimuli as a suppressor. The amount of suppression increased with the level of the harmonic complex [Fig. 20(c)], and, for a fixed level of the harmonic complex, increased in inverse proportion to the level of the near-CF tone. Psychoacoustical experiments in humans, involving the detection of a tone in the presence of a (+/−) Schroeder-phase complex, reveal that Schroeder-phase complexes with increasing instantaneous frequency (−SCHR) provide more masking than +SCHR stimuli (Smith *et al.*, 1986; Kohlrausch and Sander, 1995; Carlyon and Datta, 1997; Summers and Leek, 1998). Earlier studies (Smith *et al.*, 1986; Kohlrausch and Sander 1995) concluded that this phenomenon resulted from phase characteristics of basilar membrane motion. Using a linear model of BM motion those authors noticed that the ''gaps,'' or valleys, in BM

responses to +SCHR stimuli were wider than those seen in responses to −SCHR stimuli, thus allowing the detection of the probe signal at lower thresholds. Our results certainly are consistent with those obtained in psychophysical experiments, although the phenomenon of masking probably represents the occurrence of other processes at higher levels in the auditory system. Also, our analysis (e.g., Fig. 20) only looks at one frequency component of the response, the response to the tone, not at the overall BM waveform. Interestingly, behavioral experiments performed in birds (Leek *et al.*, 2000), akin to those in humans, have shown that +SCHR stimuli provide an amount of masking similar to, and in some cases better than, that given by −SCHR stimuli, a consequence of differences in the processing of sounds in the mammalian and avian auditory system.

The results of our experiments generally agree with the notion that a swept stimulus with rising frequency is transmitted more ''efficiently'' through the cochlea than a stimulus with falling frequency (Shore and Nuttall, 1985; Dau *et al.*, 2000). In the context of this work, efficiency can be measured by estimating the amount of energy (or rms value) in BM responses. Waveforms of responses to −SCHR contain larger rms values than those obtained in responses to +SCHR stimuli (Sec. IV A). This conclusion appears, however, contradictory with the observation that the BM represents the instantaneous frequency of the +SCHR better (i.e., with a similar slope) than that of the −SCHR (Fig. 16).

Simulation results using gamma-tone filters, which do not show any difference in ''peakiness'' in the responses to both types of stimuli used here, might provide a clue to understanding BM responses to these stimuli. The impulse response of the gamma-tone filter does not show any type of frequency modulation, which is a property of BM click responses. Group delay versus frequency curves of gamma-tone impulse responses shows a maximum at the frequency of their response peak, whereas for BM click responses the peak in group delay curves always occurs above CF. Those two properties of BM click responses are probably responsible for most of the differences between the responses to the two suppressors (−SCHR and +SCHR) observed in our experimental results. There are also differences with stimulus level, which are obviously due to the BM nonlinearity. In fact, all the differences in the responses increase with a decrease in stimulus level, showing a correlation with the gain of the cochlear amplifier.

ACKNOWLEDGMENTS

This work was supported by the National Institute of Deafness and Communications Disorders, Grant No. R01 DC01910. Thanks to Nigel P. Cooper, C. Dan Geisler, and Van Summers for their useful comments on an earlier version of this manuscript.

Allen, J. B. (**1983**). ''Magnitude and phase-frequency response to single tones in the auditory nerve,'' J. Acoust. Soc. Am. **73**, 2071–2092.

Bennett, W. R. (**1970**). *Introduction to Signal Transmission* (McGraw-Hill, New York).

Békésy, G. v. (**1949**). ''On the resonance curve and the decay period at various points on the cochlear partition,'' J. Acoust. Soc. Am. **21**, 245–254.

Carlyon, R. P., and Datta, A. J. (**1997**). ''Excitation produced by Schroeder-phase complexes: Evidence for fast-acting compression in the auditory system,'' J. Acoust. Soc. Am. **101**, 3636–3647.

Cooper, N. P., and Rhode, W. S. (**1992**). ''Basilar membrane mechanics in the hook region of cat and guinea-pig cochleae: Sharp tuning and nonlinearity in the absence of baseline position shifts,'' Hear. Res. **63**, 163–190.

Cooper, N. P. (**1996**). ''Two-tone suppression in cochlear mechanics,'' J. Acoust. Soc. Am. **99**, 3087–3098.

Cooper, N. P., and Rhode, W. S. (**1996**). ''Fast travelling waves, slow travelling waves and their interactions in experimental studies of apical cochlear mechanics,'' Aud. Neurosci. **2**, 289–299.

Dau, T., Wegner, O., Meller, V., and Kollmeier, B. (**2000**). ''Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion,'' J. Acoust. Soc. Am. **107**, 1530–1540.

de Boer, E., and Nuttall, A. L. (**1997**). ''The mechanical waveform of the basilar membrane. I. Frequency modulations ('glides') in impulse responses and cross-correlation functions,'' J. Acoust. Soc. Am. **101**, 3583–3592.

Greenwood, D. D. (**1990**). ''A cochlear frequency-position for several species—29 years later,'' J. Acoust. Soc. Am. **87**, 2592–2605.

Kochhar, R. (**1992**). ''Effective SPL computation,'' Programming note no. 6, http://www.physiology.wisc.edu/comp/docs/notes/not006.html.

Kohlrausch, A., and Sander, A. (**1995**). ''Phase effects in masking related to dispersion in the inner ear. II. Masking period patterns of short targets,'' J. Acoust. Soc. Am. **97**, 1817–1829.

Leek, M. R., Dent, M. L., and Dooling, R. (**2000**). ''Masking by harmonic complexes in budgerigars (*Melopsittacus undulatus*),'' J. Acoust. Soc. Am. **107**, 1737–1744.

Lyon, R. F. (**1997**). ''All-pole models of auditory filtering,'' in *Diversity in Auditory Mechanics*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, and E. Hecht-Poinar (World Scientific, Singapore), pp. 205–211.

Narayan, S. S., and Ruggero, M. A. (**2000**). ''Basilar-membrane mechanics at the hook region of the chinchilla cochlea,'' in *Recent Developments in Auditory Mechanics*, edited by H. Wada, T. Takasaka, K. Ikeda, K. Ohyama, and T. Koike (World Scientific, Singapore), pp. 95–101.

Nuttall, A. L., and Dolan, D. F. (**1996**). ''Steady-state sinusoidal velocity responses of the basilar membrane in guinea pig,'' J. Acoust. Soc. Am. **99**, 1556–1565.

Papoulis, A. (**1977**). *Signal Analysis* (McGraw-Hill, New York).

Recio, A., Narayan, S. S., and Ruggero, M. A. (**1997**). ''Wiener-kernel analysis of basilar-membrane responses to white noise,'' in *Diversity in Auditory Mechanics*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, and E. Hecht-Poinar (World Scientific, Singapore), pp. 325–331.

Recio, A., Rich, N. C., Narayan, S. S., and Ruggero, M. A. (**1998**). ''Basilar-membrane responses to clicks at the base of the chinchilla cochlea,'' J. Acoust. Soc. Am. **103**, 1972–1989.

Rhode, W. S. (**1971**). ''Observations of the vibration of the basilar membrane in squirrel monkey using the Mössbauer technique,'' J. Acoust. Soc. Am. **49**, 1218–1231.

Rhode, W. S., and Cooper, N. P. (**1993**). ''Two-tone suppression and distortion production on the basilar membrane in the hook region of cat and guinea pig cochleae,'' Hear. Res. **66**, 31–45.

Rhode, W. S., and Recio, A. (**2000**). ''Study of mechanical motions in the basal region of the chinchilla cochlea,'' J. Acoust. Soc. Am. **107**, 3317–3332.

Robles, L., Rhode, W. S., and Geisler, C. D. (**1976**). ''Transient response of the basilar membrane measured in squirrel monkey using the Mössbauer effect,'' J. Acoust. Soc. Am. **59**, 926–939.

Robles, L., Ruggero, M. A., and Rich, N. C. (**1986**). ''Basilar-membrane mechanics at the base of the chinchilla cochlea. I. Input-output functions, tuning curves and response phases,'' J. Acoust. Soc. Am. **80**, 1364–1374.

Robles, L., Ruggero, M. A., and Rich, N. C. (**1997**). ''Two-tone distortion on the basilar membrane of the chinchilla cochlea,'' J. Neurophysiol. **77**, 2385–2399.

Ruggero, M. A. (**1980**). ''Systematic errors in indirect estimates of basilar membrane travel times,'' J. Acoust. Soc. Am. **67**, 707–710.

Ruggero, M. A., Robles, L., and Rich, N. C. (**1992**). ''Two-tone suppression in the basilar membrane of the cochlea: Mechanical basis of auditory-nerve rate suppression,'' J. Neurophysiol. **68**, 1087–1099.

Ruggero, M. A., Robles, L., Rich, N. C., and Recio, A. (**1992**). ''Basilar membrane responses to two-tone and broadband stimuli,'' Philos. Trans. R. Soc. London, Ser. B **336**, 307–315.

Ruggero, M. A., Rich, N. C., and Recio, A. (**1996**). ''The effect of intense acoustic stimulation on basilar-membrane vibrations,'' Aud. Neurosci. **2**, 329–345.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (**1997**). ''Basilar-membrane responses to tones at the base of the chinchilla cochlea,'' J. Acoust. Soc. Am. **101**, 2151–2163.

Russell, I. J., and Nilsen, K. E. (**1997**). ''The location of the cochlear amplifier: Spatial representation of a single tone on the guinea pig basilar membrane,'' Proc. Natl. Acad. Sci. U.S.A. **94**, 2660–2664.

Schroeder, M. R. (**1970**). ''Synthesis of low peak-factor signals and binary sequences with low autocorrelations,'' IEEE Trans. Inf. Theory **16**, 85–89.

Shore, S. E., and Nuttall, A. L. (**1985**). ''High-synchrony cochlear compound action potentials evoked by rising frequency-swept tone burst,'' J. Acoust. Soc. Am. **78**, 1286–1295.

Smith, B. K., Sieben, U. K., Kohlrausch, A., and Schroeder, M. R. (**1986**). ''Phase effects in masking related to dispersion in the inner ear,'' J. Acoust. Soc. Am. **80**, 1631–1637.

Strube, H. W. (**1985**). ''A computationally efficient basilar-membrane model,'' Acustica **58**, 207–214.

Summers, V., and Leek, M. R. (**1998**). ''Masking of tones and speech by Schroeder-phase harmonic complexes in normally hearing and hearing-impaired listener,'' Hear. Res. **118**, 139–150.

Wever, E. G., and Lawrence, M. (**1971**). *Physiological Acoustics* (Princeton University Press, Princeton, NJ).

Effect of turgor pressure on outer hair cell motility

M. Adachi,[a] M. Sugawara,[b] and K. H. Iwasa[c]
Biophysics Section, Laboratory of Cellular Biology, National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Bethesda, Maryland 20892-0922

(Received 4 April 2000; accepted for publication 9 August 2000)

The motor in the outer hair cell converts electrical energy directly into mechanical energy. There are two possible mechanisms for such a motor: one depends on changes in the membrane area (''area motor'') and the other on changes in the stiffness (''stiffness motor''). These two mechanisms, which are not mutually exclusive, give different predictions on turgor pressure dependence of the amplitude. It was found that an increased pressure shifts the voltage dependence but does not change the amplitude of length changes. This observation is incompatible with a pure stiffness motor model. It was also confirmed that length changes are closely related to charge movements by monitoring the membrane capacitance. It can be concluded that hair cell motility is primarily based on area changes, and not changes in the elastic moduli. [S0001-4966(00)02811-3]

PACS numbers: 43.64.Ld, 43.64.Kc, 43.64.Bt [LHC]

I. INTRODUCTION

The outer hair cell, which has a cylindrical cell body, changes its length with the membrane potential (Brownell *et al.*, 1985; Ashmore, 1987; Santos-Sacchi and Dilger, 1988). This motility, which is sometimes referred to as electromotility, is thought to make the outer hair cell a critical factor both for fine tuning and for the wide dynamic range of the mammalian ear (Mountain, 1980; Liberman and Dodds, 1984; de Boer, 1991; Zweig, 1991). The frequency range of this motility exceeds 20 kHz, consistent with the auditory range (Dallos and Evans, 1995; Gale and Ashmore, 1997; Frank *et al.*, 1999). The direct dependence of this motility on the membrane potential is due to charges transferable across the membrane (Ashmore, 1990; Santos-Sacchi, 1991; Iwasa, 1993), analogous to gating charges of voltage gated ion channels or those of transporters. The movement of these charges is coupled with mechanical changes in the motor. This coupling is evidenced not only by the observation that length changes coincide with charge transfers, but also by the observation that externally applied tension reciprocally affects charge transfer (Iwasa, 1993; Gale and Ashmore, 1994; Kakehata and Santos-Sacchi, 1995). This mechanism enables the cell to directly convert electrical energy into mechanical energy. In general, there are two kinds of mechanical changes that are associated with such motile mechanisms. One such change is in the membrane area and the other in the stiffness (Iwasa, 1997). The motile mechanism which depends on membrane area changes of the motor can be called an ''area motor'' (Iwasa, 1993; Santos-Sacchi, 1993) or ''displacement motor'' (Dallos *et al.*, 1993) and the other a ''stiffness motor.'' It has been shown that area motors (Iwasa, 1994; Iwasa and Adachi, 1997) are successful in explaining the motility of the outer hair cell.

Among recent papers on the motility of the outer hair cell, two experimental reports (He and Dallos, 1999; Adachi and Iwasa, 1999) appear, at first sight, to contradict each other. One of them shows that the axial stiffness of the cell body is dependent on the membrane potential (He and Dallos, 1999). The other reports that the membrane motor of cells internally treated with trypsin undergoes area changes tightly coupled with charge transfer across the membrane (Adachi and Iwasa, 1999). These observations raise two questions: ''Can stiffness changes explain the electromotility?'' and ''Can the area motor predict changes in the axial stiffness?''

The second question was answered, at least in part, by a report that the area motor does give rise to voltage-dependent stiffness changes similar to those obtained experimentally (Iwasa, 2000). We intend to address the first question by experimentally examining the effect of turgor pressure on the amplitude of the motility. That is because the most discernible predictions are obtained from the two motor mechanisms on the amplitude. An area motor model predicts that the amplitude of the electromotile response of the cell is independent of pressure difference across the membrane. A stiffness motor predicts that the amplitude of the electromotility linearly increases with pressure difference.

The two models also give different predictions on shift in the voltage dependence of the cell length (and of the membrane capacitance). However, different predictions in voltage shift are more difficult to distinguish. The voltage shift expected from the area motor has a linear dependence on pressure whereas the one expected from the stiffness motor has a quadratic dependence (Iwasa, 1997). Distinguishing such a difference requires experimental data with high precision in a relatively wide pressure range.

The critical experiment for distinguishing between the two models is thus to determine how the amplitude of voltage-driven motility depends on turgor pressure.

[a]Present address: Department of Otolaryngology, Hamamatsu University School of Medicine, Handa-cho, Hamamatsu 431-3142, Japan.
[b]Present address: Department of Mechanical Engineering, Tohoku University, Aoba-ku, Sendai 980-7859, Japan.
[c]Author to whom correspondence should be addressed. Electronic mail: iwasa@nih.gov

II. THEORY: AMPLITUDE PREDICTED BY TWO MOTOR MODELS

The outer hair cell can be approximated by a cylinder of radius r, which is supported by its elastic lateral wall. An area motor and a stiffness motor are two motile mechanisms that can change the length of such a cell. Here, we examine the amplitude of length changes predicted by the two motile mechanisms. The voltage dependence of the capacitance and that of the cell length has been presented already (Iwasa, 1997). The previous treatment of the stiffness motor was limited to a special case in which mechanical isotropy of the lateral membrane is assumed. The present treatment extends the validity to include mechanical orthotropy.

To describe the lateral membrane of a cylindrical cell, we define the axial strain ϵ_z and the circumferential strain ϵ_c as follows:

$$\epsilon_z = \frac{L - L_0}{L_0}, \quad \epsilon_c = \frac{r - r_0}{r_0}.$$

Here, the axial length L and the radius r of the cell are L_0 and r_0, respectively, at stress-free reference condition. We initially assume that the strains are entirely elastic. Thus, $\epsilon_z = \epsilon_{ez}$ and $\epsilon_c = \epsilon_{ec}$, where the suffix e signifies an elastic component. We assume that the strains are small.

Since elastic force is balanced by the force due to pressure difference P across the cell membrane and an external axial force, we obtain a set of constitutive equations for tension (Iwasa and Adachi, 1997)

$$d_z \epsilon_{ez} + c\,\epsilon_{ec} = \frac{r_0 P}{2} + T_{xt}, \tag{1}$$

$$c\,\epsilon_{ez} + d_c \epsilon_{ec} = r_0 P. \tag{2}$$

Here, the term T_{xt} is tension due to an external axial force F_{xt} ($F_{xt} = 2\pi r_0 T_{xt}$). The quantities d_z, d_c, and c are elastic moduli of the lateral membrane, which is assumed orthotropic (Tolomeo and Steele, 1995; Iwasa and Adachi, 1997). If the membrane is isotropic, $d_z = d_c = K + \mu$ and $c = K - \mu$, where K is the area modulus and μ the shear modulus. The diagonal elements d_z and d_c and the off-diagonal element c satisfy an inequality $d_z d_c > c^2$. Previously obtained values for these moduli are $d_z = 43$ mN/m, $d_c = 61$ mN/m, and $c = 42$ mN/m (Iwasa and Adachi, 1997).

Because of the cylindrical geometry of the cell, the volume strain ϵ_v is related to length strains as $\epsilon_v = \epsilon_z + 2\epsilon_c$, which is proportional to pressure P in the load-free condition. In other words, the volume strain is proportional to turgor pressure.

A. Area motor model

If the motor functions by changing the surface area, the total axial strain ϵ_z is expressed by $\epsilon_z = \epsilon_{ez} + \epsilon_{mz}(V)$, where ϵ_{ez} is the elastic strain and $\epsilon_{mz}(V)$ is strain due to the motor, which is dependent on the membrane potential V. The assumption is that changes in the motor strains drive the cell. An analogous relationship is written for the circumferential strains.

The simplest example for such a motor is a membrane molecule with two states that differ both in the membrane area and in surface charge. Due to the difference in surface charge in the two states, a change in the membrane potential favors one state over the other, leading to a change in the membrane area of the cell. Since such a molecule has both electric energy and mechanical energy, a change in membrane tension affects the energy of the motor. Thus such a motile mechanism has a shift in the voltage dependence proportional to the change in tension (Iwasa, 1993, 1994; Iwasa and Adachi, 1997).

By eliminating the elastic strains in Eqs. (1) and (2), we obtain

$$d_z \epsilon_z + c\,\epsilon_c - [d_z \epsilon_{mz}(V) + c\,\epsilon_{mc}(V)] = \frac{r_0 P}{2} + T_{xt}, \tag{3}$$

$$c\,\epsilon_z + d_c \epsilon_c - [c\,\epsilon_{mz}(V) + d_c \epsilon_{mc}(V)] = r_0 P. \tag{4}$$

Because the motor activity can take place in a short duration of time, it will be reasonable to assume that the cell volume is constant. Thus, we can write

$$\epsilon_z + 2\epsilon_c = \epsilon_v, \tag{5}$$

where ϵ_v is the volume strain, which remains constant during fast voltage changes.

The volume strain ϵ_v can be changed by changing pressure P applied to the cell. A pressure increase is realized, for example, by fluid injection and by osmotic means. If such pressure application is done at the holding potential in a voltage range in which the motor is not very voltage sensitive, the volume strain ϵ_v is approximately proportional to pressure P, which can be called turgor pressure. It is approximately proportional because the strains include motor terms which do not increase linearly with pressure. The effect of turgor pressure is represented by the volume strain ϵ_v.

By combining Eqs. (3), (4), and (5), we obtain

$$(4d_z - 4c + d_c)\epsilon_z + (2c - d_c)\epsilon_v = 2(2d_z - c)\epsilon_{mz}(V) + 2(2c - d_c)\epsilon_{mc}(V) + 4T_{xt}. \tag{6}$$

In Eq. (6), strains $\epsilon_{mz}(V)$ and $\epsilon_{mc}(V)$ are motor terms, which are combination of the elastic moduli and the motor strains. The term that involves the volume strain ϵ_v remains constant and thus it does not contribute to the amplitude. Namely, for constant T_{xt}, which include the load-free condition, the amplitude of the axial strain is given by

$$\epsilon_z(-\infty) - \epsilon_z(\infty) = \frac{2(2d_z - c)\epsilon_{mz}(-\infty) + 2(2c - d_c)\epsilon_{mc}(-\infty)}{4d_z - 4c + d_c}. \tag{7}$$

To derive the amplitude, we assumed that $\epsilon_{mz}(\infty) = \epsilon_{mc}(\infty) = 0$. The motor strains in both extremes of the voltage do not depend on pressure or tension. The amplitude is independent of the volume strain and thus independent of turgor pressure.

The axial stiffness of the cell, which can be expressed by $2\pi r T_z / \epsilon_z$ is less than the value of $\frac{1}{2}(4d_z - 4c + d_c)\pi r$ in the voltage range where the motor state is sensitive to tension (Iwasa, 2000).

B. Stiffness motor model

Next, we examine how changes in the elastic moduli alone affect the cell length. We assume that the strains are all elastic and that the elastic moduli d_z, d_c, and c depend on the membrane potential V. Because the action of the motor is through changing the stiffness moduli and not through changing the membrane area directly, motor strains are absent. Thus, the elastic strains in Eqs. (1) and (2) are the same as the total strains ϵ_{ez} and ϵ_{ec}. Namely

$$d_z(V)\epsilon_z + c(V)\epsilon_c = \frac{r_0 P}{2} + T_{xt},$$

$$c(V)\epsilon_z + d_c(V)\epsilon_c = r_0 P.$$

We also assume that the volume constant condition (5) holds. We then obtain

$$\epsilon_z = -\epsilon_v \cdot \frac{2c(V) - d_c(V)}{4d_z(V) - 4c(V) + d_c(V)} + \frac{4T_{xt}}{4d_z(V) - 4c(V) + d_c(V)}. \quad (8)$$

Since the second term is the response to external force F_{xt} $(=2\pi r_0 T_{xt})$, the axial stiffness of the cell is $\frac{1}{2}\pi r_0[4d_z(V) - 4c(V) + d_c(V)]$.

In the load-free condition $(T_{xt}=0)$, the axial strain changes if the stiffness ratio $[2c(V) - d_c(V)]/[4d_z(V) - 4c(V) + d_c(V)]$ depends on the membrane potential. The amplitude also depends on the volume strain ϵ_v, which is related to pressure P at the holding potential V_h, i.e., turgor pressure P_{trg}, with

$$\epsilon_v = \frac{r_0 P_{trg}}{2} \frac{4d_z(V_h) - 4c(V_h) + d_c(V_h)}{d_z(V_h)d_c(V_h) - c^2(V_h)}.$$

Whether the length increases or decreases in response to changes in the membrane potential in the load-free condition is determined by changes in the ratio of the elastic moduli and not by changes in the axial stiffness. To illustrate the point, we set $A(V) = 2(2d_z(V) - c(V))$ and $B(V) = 2c(V) - d_c(V)$. The axial strain in the load-free condition is given by,

$$\epsilon_{z\ \text{load-free}} = -\epsilon_v \cdot \frac{A(V)}{A(V) - B(V)}, \quad (9)$$

while the axial stiffness is proportional to $A(V) - B(V)$. The load-free strain changes if the ratio $A(V)/B(V)$ depends on the membrane potential. It does not depend on whether the axial stiffness, which is proportional to $A(V) - B(V)$, increases or decreases. For a given change of the ratio A/B, the load-free amplitude is proportional to ϵ_v and thus it is proportional to turgor pressure P_{trg}.

It is intuitively obvious that the axial stiffness is proportional to a linear combination of elastic moduli. It is not obvious that a ratio of the elastic moduli determines load-free length changes. Two factors contribute to this result. One is that the membrane is stretched by anisotropic tension, both components of which are proportional to pressure. The other is the volume constant condition. As the consequence of the combination, the resulting membrane deformation is determined by the ratio of the effects of the tension in thetwo directions. Because the elastic moduli determine the effectiveness of tension in inducing displacements, a ratio between elastic moduli determines the deformation. This is illustrated by a special case of isotropic membrane $(d_z = d_c = K + \mu$ and $c = K - \mu)$. Substituting d_z, d_c, and c in Eq. (8) with the area modulus K and the shear modulus μ, we obtain

$$\epsilon_{z\ \text{load-free}} = -\epsilon_v \cdot (1 - 3\mu/K)/(1 + 9\mu/K).$$

The ratio μ/K has a one-to-one relationship with Poisson's ratio ν, which is defined by $(K - \mu)/(k + \mu)$ (Fischer, 1992). Thus, the pressure deformation of a cell with isotropic membrane is controlled by Poisson's ratio.

In this motile mechanism, a change in pressure shifts the voltage dependence proportionally to the square of turgor pressure (Iwasa, 1997). This is because the elastic energy of the unit area of the lateral cell membrane is proportional to the square of the strain, and the strain is proportional to pressure.

III. METHODS

A. Cell preparation

Isolated outer hair cells were prepared from pigmented guinea pigs with the method previously described (Iwasa, 1993; Adachi and Iwasa, 1999). The length of the cells used for the experiment ranged between 40 and 75 μm. Isolated cells were chosen for the experiment. In some of the experiments, dispase (Boehringer-Mannheim) treatment (0.5 units/ml for 15 min at 21 °C) was used before mechanical isolation to increase the yield of isolated cells. The animal procedures were performed in accordance with a protocol approved by NINDS/NIDCD (animal protocol 902-99).

B. Patch clamping

The resistance of pipettes used was between 1.5 and 2.5 MΩ when filled with an intracellular medium. The cells which showed a zero-current potential more negative than −50 mV were used for the experiment using nonblocking media. When the blocking media were used, the zero-current potential differed little from 0 mV so that the zero-current potential did not serve as an easy test for the seal. The access resistance was between 4 and 6 MΩ. The membrane resistance of the cells used was between 40 and 100 MΩ with the nonblocking media and about 300–600 MΩ with the blocking media near the holding potential of −75 mV. The experiments were performed at room temperature (22±1)°C.

A patch amplifier (APC-8, Medical Systems, Greenvale, NY) was used for whole-cell voltage clamp experiments. Voltage pulses were generated with an ITC-18 interface (Instrutech, Great Neck, NY) using the IGOR program (Wave Metrics, Lake Oswego, OR) with a software module created by Richard J. Bookman's laboratory at the University of Miami School of Medicine (http: //chroma.med.miami.edu/cap/). The current output was filtered with a 10-kHz Bessel filter.

C. Media, hypo-osmotic perfusion, and pressure application

The nonblocking internal medium contained 145 mM KCl, 2 mM $MgCl_2$, 1 mM EGTA, 0.555 mM $CaCl_2$, and 10 mM HEPES. The nonblocking external medium contained 135 mM NaCl, 5 mM KCl, 2 mM $MgCl_2$, 1.5 mM $CaCl_2$, 5 mM glucose, and 10 mM HEPES. The *p*H of both media was adjusted to 7.4. Osmolarity was adjusted to 300 mOsm/kg with glucose.

Ion channel blocking solutions were used for capacitance experiments. The internal blocking solution consisted of 140 mM CsCl, 2 mM $CaCl_2$, 5 mM EGTA, and 10 mM HEPES. The external blocking solution contained 140 mM NaCl, 5 mM CsCl, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 2 mM $CoCl_2$ and 10 mM HEPES. The osmolarity of both media was adjusted to 300 mOsm/kg with glucose. In all media the *p*H was adjusted to 7.4.

The hypo-osmotic media used were made by mixing the equiosmolar media with respective Na^+-less external media. The Na^+-less media, both blocking and nonblocking, were the same as respective equiosmolar media except that they did not contain Na^+. Usually 2:1 mixtures of regular osmolar media and Na^+-less media were used. The perfusion pipettes were made of the same capillary but their tip diameters were about 5 μm. The pressure applied to the perfusion pipette was controlled with a three-way solenoid valve. One of the connections was to an air reservoir with about 1 kPa pressure produced by a small adjustable air pump. The third connection was closed.

Positive pressure was applied to the pipette through a solenoid valve (General Valve, Fairfield, NJ), which was connected to a reservoir with a diaphragm air pump (Thomas Scientific, Swedesboro, NJ) driven with a variable power supply. The pressure at the reservoir was monitored with a pressure gauge (model 68370, Cole-Parmer, Vernon Hills, IL).

D. Voltage waveform and determination of the membrane capacitance

We used a staircase voltage waveform to elicit length changes of the cell as well as to measure the membrane capacitance. The usual waveform consisted of a jump from the holding potential of -75 to -150 mV, 15 mV steps from -150 to 105 mV, -15 mV steps down to -150 mV, and a jump back to the holding potential. The duration of the waveform was 1.2 s. In experiments in which pipette pressure was applied, the upper limit of the waveform was decreased to keep the seal. It was achieved by reducing the number of voltage steps. In those cases the total duration of the waveforms was shorter.

The membrane capacitance was determined from current responses to voltage steps assuming an equivalent circuit that consists of access resistance at the tip of the patch pipette, the membrane resistance, and the membrane capacitance of the cell (Gillis, 1995). We used the standard relationships between relaxation time τ of the transient current, the access resistance R_a, the membrane resistance R_m, and the membrane capacitance C_m, the charge ΔQ transferred by a voltage step ΔV, and a change of steady-state current ΔI_s before and after the voltage step. Namely

$$\frac{1}{\tau}=\left(\frac{1}{R_a}+\frac{1}{R_m}\right)\frac{1}{C_m},$$

$$\Delta Q=\Delta V C_m\cdot\frac{R_m}{R_a+R_m},$$

$$\Delta I_s=\frac{\Delta V}{R_a+R_m}.$$

Because the capacitive current that contributes to ΔQ is an exponential function of time, it is convenient to replace the second equation for ΔQ by

$$\Delta I_{\max}=\frac{\Delta V}{R_a},$$

where $\Delta I_{\max}$ is the peak capacitive current extrapolated to time zero when current through the resister R_m is not increased. From experimental values for τ, $\Delta I_{\max}$, and ΔI_s, the membrane capacitance, the membrane resistance, and the access resistance were determined. The method was tested using dummy cells.

E. Determination of cell strain

Images of the cell were stored with a video recorder and digitized off-line with an image grabber card (Scion, Frederick, MD) using a computer program (The NIH IMAGE, W. Rasband, NIMH). The accuracy in determining the axial strain of the cell was enhanced by fitting the light intensity peaks which corresponded to cell edges with a parabola. The number of points used for the fit was usually five. The process was automated with a macro for the NIH IMAGE program as previously described (Iwasa and Adachi, 1997).

IV. RESULTS

A. Direct pressure application through pipette

First, we applied known pressure through the recording patch pipette and applied a staircase waveform to monitor the amplitude of the voltage-induced length changes [Fig. 1(A)].

We found that increased pressure introduces a rocking mode of motion relative to the pipette in the voltage-dependent cell motion. The rocking motion complicates the measurement of cell length. In addition, the membrane potential could not be changed beyond a relatively narrow range to maintain the seal (Fig. 1). The range in the membrane potential limited the accuracy in determining the amplitude of length changes. Such problems did not occur for hypo-osmotic perfusion (see the next section). For these reasons, we decided to use hypo-osmotic perfusion as a means of applying pressure. To evaluate applied pressure during osmotic experiments, we needed to use indicators of pressure. One such indicator can be the axial strain at the holding potential. However, because the cell membrane may undergo mechanical relaxation (Ehrenstein and Iwasa, 1996), the axial strain may not be reliable for a prolonged exposure to hypo-osmotic media. Another pressure indicator is the mid-

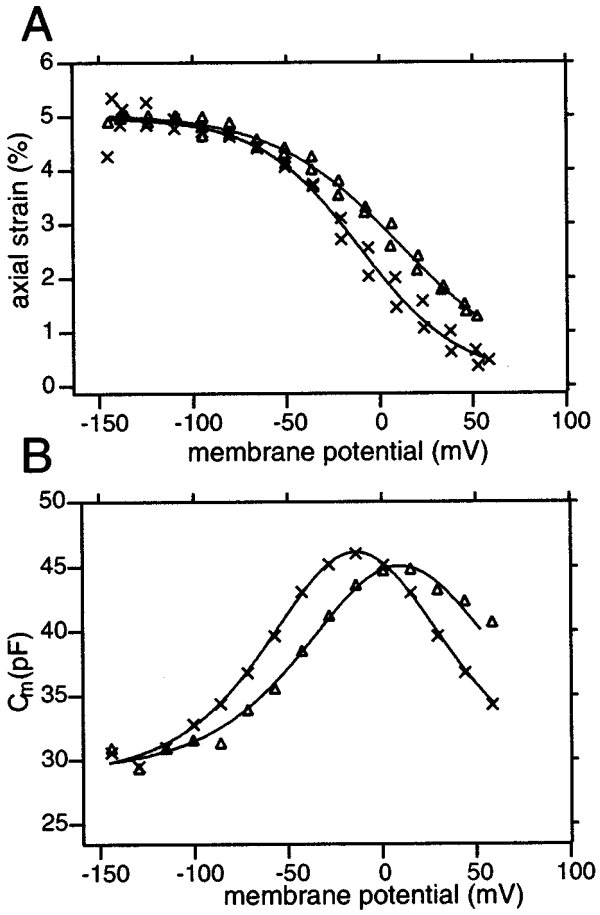


FIG. 1. Effect of pressure on the voltage dependence of the length and membrane capacitance. The cell is in the whole-cell recording configuration. Pipette pressure is 0 kPa (×) and 0.4 kPa (△). The holding potential is −75 mV. (A) The axial strain $(L(V)-L_0)/L_h$ due to changes in the membrane potential. The axial strain is normalized to the cell length L_h at the holding potential at zero pipette pressure. The values for the parameters are: $V_{mp}=(-10.6\pm3.9)$ mV, the apparent unit motor charge $q_l=(0.94\pm0.12)e$, and $\Delta L_{max}/L_h=(4.9\pm0.3)\%$ (×). $V_{mp}=(11.6\pm6.9)$ mV, $q_l=(0.74\pm0.08)e$, and $\Delta L_{max}/L_h=(4.9\pm0.5)\%$ (△). (B) The membrane capacitance plotted against the membrane potential. The values for the parameters are: $C_{lin}=(28.8\pm0.2)$ pF, $q_c=(0.81\pm0.02)\,e$, $Q=(2.2\pm0.1)$ pC, $V_{pk}=(-13.7\pm0.4)$ mV (×); $C_{lin}=(29.1\pm0.3)$ pF, $q_c=(0.75\pm0.03)\,e$, $Q=(2.1\pm0.1)$ pC, $V_{pk}=(9.4\pm1.0)$ mV (△).

point voltage of the length change or the peak voltage of the capacitance. The tension dependence of the membrane capacitance (Iwasa, 1993; Gale and Ashmore, 1994; Kakehata and Santos-Sacchi, 1995) provides a basis for such use.

During pressure application from the pipette, the voltage dependence of the membrane capacitance was reliably monitored [Fig. 1(B)]. The membrane capacitance C_m has a bell shaped dependence on the membrane potential V. This voltage dependence is given by (Ashmore, 1990; Santos-Sacchi, 1991; Iwasa, 1993)

$$C_m(V)=C_{\rm lin}+\frac{q_cQ\exp[q_c(V-V_{\rm pk})/k_BT]}{k_BT(1+\exp[q_c(V-V_{\rm pk})/k_BT])^2}, \quad (10)$$

which is obtained taking the voltage derivative of a two state Boltzmann function. Here, $C_{\rm lin}$ is the regular (linear) membrane capacitance, q_c is the effective unit charge of the motor for the capacitance, Q is the total charge of the motor in the cell, $V_{\rm pk}$ is the peak voltage of the capacitance, k_B is Boltzmann's constant, and T is the temperature.

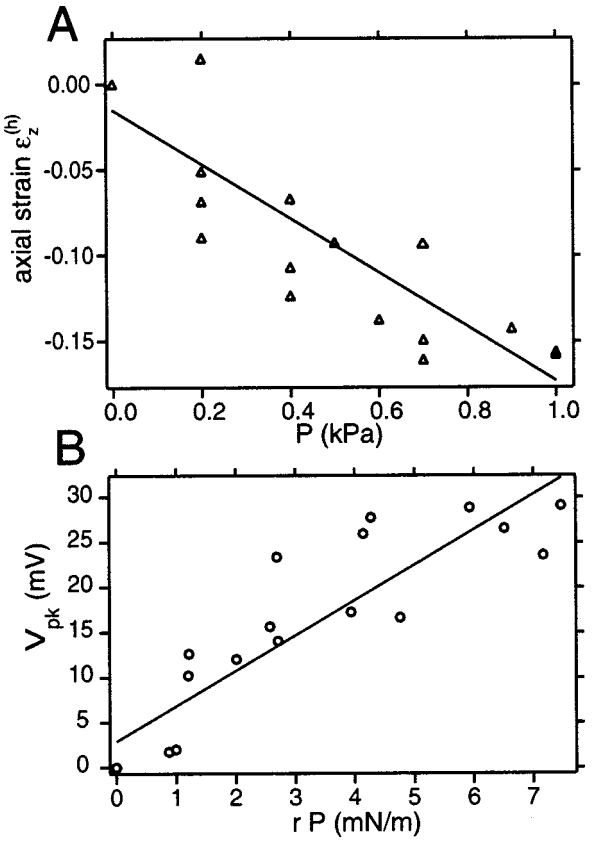


FIG. 2. Effect of pressure on the axial strain $\epsilon_z^{(h)}$ at the holding potential and on the peak potential $V_{\rm pk}$ of the capacitance. The holding potential V_h is −75 mV. Data points from four cells are pooled. (A) The axial strain $\epsilon_z^{(h)}$ at the holding potential is plotted against pressure applied at the recording pipette. The strain is normalized to the cell length at the holding potential while pipette pressure is zero. The slope is (6.7 ± 0.5) kPa per unit axial strain. (B) The shift of the peak potential $V_{\rm pk}$ of the capacitance plotted against tension rP. The slope is (3.9 ± 0.5) mV per 1 mN/m. Tension in a cylindrical cell is not isotropic. The quantity rP corresponds to the circumferential tension. Tension in the axial direction is $\frac{1}{2}rP$. If the peak shift is plotted against pressure, the slope is (27.5 ± 3.1) mV/kPa.

We found that the cell length at the holding potential of −75 mV correlated well with pressure applied from the pipette [Fig. 2(A)]. We also found that the shift of the potential that maximizes the capacitance shows a good correlation with membrane tension at the holding potential [Fig. 2(B)]. Thus, we can use these plots to evaluate membrane tension and pressure during osmotic experiments.

B. Pressure application by hypo-osmotic perfusion

We first examined the relationship between the midpoint voltage of the length change and the capacitance peak voltage during hypo-osmotic perfusion.

The length change ΔL has a sigmoidal dependence of the membrane potential V. It is adequately described by a two-state Boltzmann function. Namely, the length L of the cell is expressed (Ashmore, 1987; Santos-Sacchi, 1991; Iwasa, 1994)

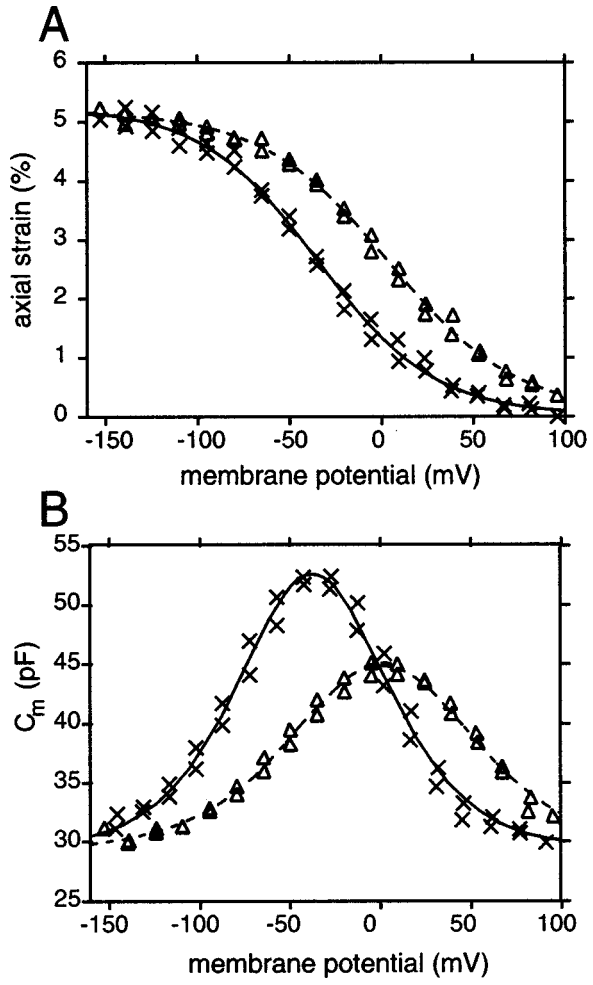


FIG. 3. An example of the voltage dependence of the membrane capacitance and the cell length before and during hypo-osmotic perfusion. (A) The axial cell strain $(L(V)-L_0)/L_h$ elicited by changes in the membrane potential before (×) and during hypo-osmotic perfusion (△). The axial strain is normalized to the cell length L_h at the holding potential before perfusion. During the perfusion, the midpoint voltage V_{mp} shifts in a positive direction, and the sharpness of the transition reduces. The difference in length due to pressure in the two conditions is disregarded for comparison. $V_{mp}=(-34.5\pm1.8)$ mV, the apparent unit motor charge $q_l=(0.82\pm0.05)\,e$, and $\Delta L_{max}/L_h=(5.2\pm0.1)\%$ (×). $V_{mp}=(4.3\pm1.9)$ mV, $q_l=(0.75\pm0.04)\,e$, and $\Delta L_{max}/L_h=(5.1\pm0.1)\%$ (△). (B) The membrane capacitance C_m before (×) and during hypo-osmotic perfusion (△). Hypo-osmotic perfusion shifts the peak to a more positive potential. The shift is accompanied by a reduction of peak height and some broadening of the peak. The peak potential $V_{pk}=(-37.6\pm1.0)$ mV, the apparent unit motor charge $q_c=(0.92\pm0.06)\,e$, the total motor charge $Q=(2.9\pm0.2)$ pC, and the linear component of the membrane capacitance $C_{lin}=(29.4\pm1.0)$ pF (×). $V_{pk}=(0.5\pm0.9)$ mV, $q_c=(0.76\pm0.03)\,e$, $Q=(2.1\pm0.1)$ pC, and $C_{lin}=(29.3\pm0.4)$ pF (△). The turgor pressure estimated from the voltage shift using the calibration shown in Fig. 2 is about 1.3 kPa.

$$L(V)=L_0+\frac{\Delta L_{max}}{1+\exp[q_l(V-V_{mp})/k_BT]}, \tag{11}$$

where L_0 is the asymptotic length of the cell at extremely positive potentials and q_l is the effective unit charge of the motor for the length changes. The model requires that q_l must be equal to q_c.

By fitting the data points with the two-state Boltzmann function (11), the midpoint potential V_{mp} and the amplitude ΔL_{max} were determined [Fig. 3(A)]. The peak potential V_{pk} of the capacitance was also determined by curve fit with Eq. (10) [Fig. 3(B)].

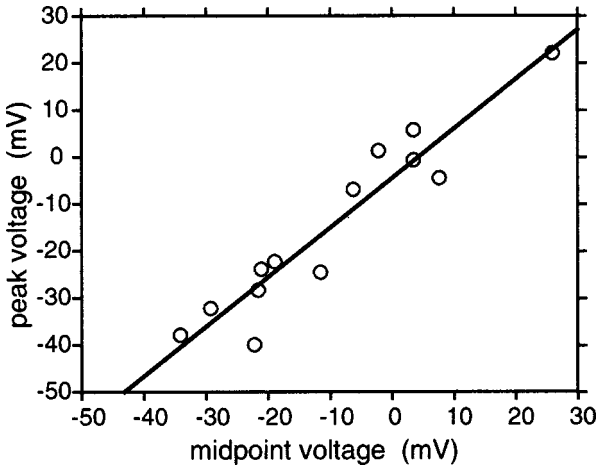


FIG. 4. The relationship between the voltage V_{pk} of the capacitance peak and the midpoint potential V_{mp} of length changes. Data points from six cells are pooled. The straight line indicates the least-squares fit to the data. The relationship is given by $V_{pk}=(1.1\pm0.1)V_{mp}-(4.5\pm2.0)$ in mV unit.

The unit charge of the motor obtained from the capacitance and length changes was similar. The value for the charge obtained from the capacitance was $(0.77\pm0.05)\,e$ and the one from the length was $(0.80\pm0.06)\,e$.

Although the peak voltage V_{pk} of the capacitance did not completely agree with the midpoint potential V_{mp} of length changes, the difference was relatively small. The midpoint potential V_{mp} was about 5 mV less than the peak voltage V_{pk} (Fig. 4). Thus, the midpoint potential V_{mp} can serve as a pressure indicator.

Now, we compare the amplitude of length changes in the load-free condition during hypo-osmotic perfusion. Pressure and membrane tension are evaluated from the indicators, namely the axial strain of the cell at the holding potential of −75 mV and the midpoint potential of length changes. We found that the amplitude of voltage dependent length changes remained constant, while pressure and tension changed as shown by their indicators (Fig. 5). According to our calibration (Fig. 2), the range of the pressure P applied was up to about 0.6 kPa and circumferential membrane tension T_c up to about 3 mN/m (these two indicators of pressure satisfy the relationship $T_c=rP$, where r is the radius of the cell which is approximately 5 μm).

V. DISCUSSION

We found that an increase in pressure difference up to 0.6 kPa across the cell membrane does not affect the amplitude of the voltage- dependent length changes. This result is consistent with the area motor model and it is inconsistent with the stiffness motor model.

All previous reports on the effect of turgor pressure on the amplitude of motility involved comparing the amplitude obtained at positive pressure and the ones obtained at negative pressure (Santos-Sacchi, 1991; Kakehata and Santos-Sacchi, 1995). While such comparisons are indeed useful to demonstrate pressure dependence, they are not useful to

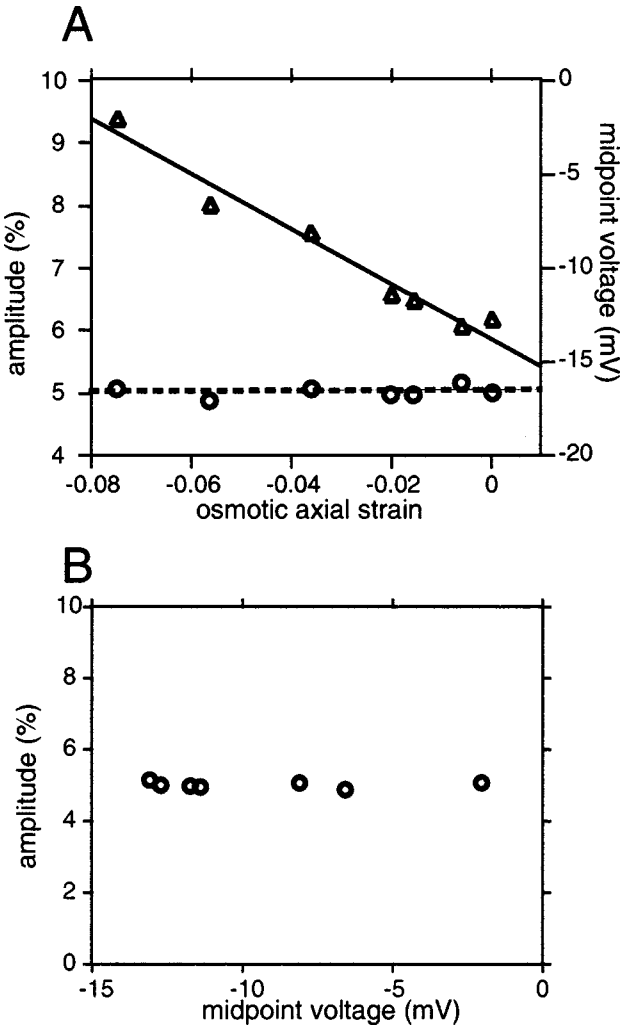


FIG. 5. The load-free amplitude $\epsilon_z^{(amp)}$ ($=\Delta L_{max}/L_h$) and the midpoint potential V_{mp} in the nonblocking media. The amplitude of the axial strain is normalized to the length L_h of the cell at the holding potential of -75 mV. Data points from five cells are pooled. (A) The amplitude $\epsilon_z^{(amp)}$ (○) and the midpoint potential V_{mp} (△) are plotted against the axial strain due to osmotic stress. The slopes of the least-mean square fits are 0.6 ± 1.4 per unit osmotic strain for the amplitude and (-144 ± 12) mV per unit osmotic strain for the midpoint potential. (B) The relationship between the midpoint potential V_{mp} and the amplitude $\epsilon_z^{(amp)}$ in the load-free condition.

clarify the motile mechanism because theoretical predictions are unavailable for a cell with negative pressure.

To describe a cell that is subjected to negative pressure, we need to introduce an additional term, which possibly involves a bending modulus, into the equations. Such a term is required to counter negative pressure. We could, however, crudely expect the behavior of a collapsed cell if pressure is near zero. In our experience, outer hair cells which have collapsed with negative pressure remain collapsed until positive pressure is applied. In such a condition, length changes decrease but still persist (unpublished observation). For null pressure, the stiffness motor predicts null length changes. We expect that such a condition reduces the efficiency of the area motor model as well. That is because area changes due to the area motor can barely affect pressure in deflated cells. A decrease in the membrane area due to the area motor results in shortening of the cell by itself. In addition, it results in an increase in cell pressure, which in turn reduces the cell length further (Iwasa and Adachi, 1997). For this reason, if pressure is not sensitive to the membrane area, the motor is less efficiente. Thus, a deflated cell is not a system useful for distinguishing the motile mechanisms.

Voltage-dependent changes in the membrane area and those in the membrane stiffness are not mutually exclusive as motor mechanisms. It would be natural to assume that states of the motor differ both in their membrane area and stiffness. We attempted to clarify the relative importance of the two factors by assuming the two pure (or extreme) cases. Our conclusion from the present experimental results is that area changes have a dominant effect in the load-free condition. Furthermore, the area motor mechanism is successful in explaining other features of outer hair cell motility as well. It provides a reasonable explanation of voltage-dependent force production (Iwasa and Adachi, 1997). The area motor model also predicts the voltage dependence of the axial stiffness, although not as large as observed. It predicts that the motor significantly reduces the axial stiffness in the voltage range where the motor is sensitive to tension (Iwasa, 2000). The effect is expected to have a maximum at about -20 mV, at which potential the membrane capacitance peaks. This mechanism is similar to the well-known displacement-dependent stiffness of the stereocilia, where gating of the transducer channel reduces the bundle stiffness (Howard and Hudspeth, 1988).

Although we have excluded stiffness changes as the motor mechanism of the outer hair cell, we did not exclude the possibility that the elastic moduli can depend on the membrane potential. To illustrate the point, let us examine a simpler case of isotropic membrane with only two elastic moduli. If the elastic moduli K and μ decrease with their ratio kept constant on depolarization, length changes in the load-free amplitude will remain unaffected but the axial stiffness will decrease. Thus, such changes in the stiffness moduli could explain the experimentally observed changes in the axial stiffness (He and Dallos, 1999), which are larger than those predicted by a pure area motor model (Iwasa, 2000).

We have assumed that the shape of the outer hair cell is supported by its lateral membrane. However, there may be an additional parallel elastic element in the core (Evans and Harris, 1995). Such an elastic element contributes to the stress in the axial direction proportional to the axial strain ϵ_z. Thus by redefining the elastic modulus d_z in Eq. (1), such an additional element can be included in the equations for the stiffness motor model. That means that our conclusions should hold even in the presence of a parallel elastic element, which is not explicitly considered.

The stress–strain relationship obtained in this experiment shows that the axial strain 0.1 corresponds to (0.67 ± 0.05) kPa, which is consistent with the value of (0.7 ± 0.3) kPa reported earlier (Iwasa and Chadwick, 1992). This agreement is somewhat surprising because a relaxation effect is expected during a prolonged pressure application (Ehrenstein and Iwasa, 1996). In the present experiment, the duration of the pressure application was up to 2 min to take a number of data points from a cell. In the previous experi-

ment, the duration of pressure application is about 10 s and only one data point was taken from a cell.

The shift of the voltage dependence of the capacitance peak is (3.9 ± 0.5) mV per 1mN/m tension, which corresponds to (27.5 ± 3.1) mV per 1 kPa pressure [Fig. 1(A)]. This value gives the membrane area change $(\Delta a_z+2\Delta a_c)$ per motor unit as about 2 nm^2, where Δa_z and Δa_c are the area difference in the two states in the axial direction and in the circumferential direction, respectively (Iwasa, 1994). The value is consistent with previous estimates (Iwasa, 1993, 1994). The observed pressure sensitivity is consistent with the about 25-mV shift for 1 kPa reported by Kakehata and Santos-Sacchi (1995).

VI. CONCLUSIONS

We can conclude that our experimental observation is consistent with an area motor and inconsistent with a stiffness motor. We analyzed our data partially based on two-state models because they are the simplest. The validity of the conclusions, however, obviously extends to more complex models. Details of the mechanism, such as the number of the states of the motor and the molecular organization of the motor, require further clarification.

Adachi, M., and Iwasa, K. H. (**1999**). ''Electrically driven motor in the outer hair cell: Effect of a mechanical constraint,'' Proc. Natl. Acad. Sci. U.S.A. **96**, 7244–7249.

Ashmore, J. F. (**1987**). ''A fast motile response in guinea-pig outer hair cells: The molecular basis of the cochlear amplifier,'' J. Physiol. (London) **388**, 323–347.

Ashmore, J. F. (**1990**). ''Forward and reverse transduction in guinea-pig outer hair cells: The cellular basis of the cochlear amplifier,'' Neurosci. Res. Suppl. **12**, S39–S50.

Brownell, W., Bader, C., Bertrand, D., and Ribaupierre, Y. (**1985**). ''Evoked mechanical responses of isolated outer hair cells,'' Science **227**, 194–196.

Dallos, P., and Evans, B. (**1995**). ''High-frequency motility of outer hair cells and the cochlear amplifier,'' Science **267**, 2006–2009.

Dallos, P., Hallworth, R., and Evans, B. N. (**1993**). ''Theory of electrically driven shape changes of cochlear outer hair cells,'' J. Neurophysiol. **70**, 299–323.

de Boer, E. (**1991**). ''Auditory physics. Physical principles in hearing theory III,'' Phys. Rep. **203**, 125–231.

Ehrenstein, D., and Iwasa, K. H. (**1996**). ''Viscoelastic ralaxation in the membrane of the auditory outer hair cell,'' Biophys. J. **71**, 1087–1094.

Evans, N. B., and Harris, D. M. (**1995**). ''Subcellular selective photoablation in isolated hair cells.'' Assoc. Res. Otolaryngol. Midwinter Meeting, **18**, p. 155.

Fischer, T. M. (**1992**). ''Is the surface area of the red blood cell membrane skeleton locally conserved?'' Biophys. J. **61**, 298–305.

Frank, G., Hemmert, W., and Gummer, A. W. (**1999**). ''Limiting dynamics of high-frequency electromechanical transduction of outer hair cells,'' Proc. Natl. Acad. Sci. U.S.A. **96**, 4420–4425.

Gale, J. E., and Ashmore, J. F. (**1994**). ''Charge displacement induced by rapid stretch in the basolateral membrane of the guinea-pig outer hair cell,'' Proc. R. Soc. London, Ser. B Biol. Sci. **255**, 233–249.

Gale, J. E., and Ashmore, J. F. (**1997**). ''An intrinsic frquency limit to the cochlear amplifier,'' Nature (London) **389**, 63–66.

Gillis, K. D. (**1995**). ''Techniques for membrane capacitance measurements,'' in *Single Channel Recording*, edited by B. Sackmann and E. Neher (Plenum, New York).

He, D. Z. Z., and Dallos, P. (**1999**). ''Somatic stiffness of cochlear outer hair cells is voltage-dependent,'' Proc. Natl. Acad. Sci. U.S.A. **96**, 8223–8228.

Howard, J., and Hudspeth, A. J. (**1988**). ''Compliance of the hair bundle associated with gating of mechanoelectrical transduction channels in the bullfrog's saccular hair cell,'' Neuron **1**, 189–199.

Iwasa, K. H. (**1993**). ''Effect of stress on the membrane capacitance of the auditory outer hair cell,'' Biophys. J. **65**, 492–498.

Iwasa, K. H. (**1994**). ''A membrane model for the fast motility of the outer hair cell,'' J. Acoust. Soc. Am. **96**, 2216–2224.

Iwasa, K. H. (**1997**). ''Mechanisms for the fast motility of the outer hair cell from the cochlea.'' in *Diversity in Auditory Mechanics*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, and E. Hecht-Poinar (World Scientific, Singapore), pp. 580–586.

Iwasa, K. H. (**2000**). ''Effect of membrane motor on the axial stiffness of the cochlear outer hair cell,'' J. Acoust. Soc. Am. **107**, 2764–2766.

Iwasa, K. H., and Adachi, M. (**1997**). ''Force generation in the outer hair cell of the cochlea,'' Biophys. J. **73**, 546–555.

Iwasa, K. H., and Chadwick, R. S. (**1992**). ''Elasticity and active force generation of cochlear outer hair cells,'' J. Acoust. Soc. Am. **92**, 3169–3173.

Kakehata, S., and Santos-Sacchi, J. (**1995**). ''Membrane tension directly shifts voltage dependence of outer hair cell motility and associated gating charge,'' Biophys. J. **68**, 2190–2197.

Liberman, M. C., and Dodds, L. W. (**1984**). ''Single neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves,'' Hear. Res. **16**, 55–74.

Mountain, D. C. (**1980**). ''Changes in endolymphatic potential and crossed olivocochlear bundle stimulation alter cochlear mechanics,'' Science **210**, 71–72.

Santos-Sacchi, J. (**1991**). ''Reversible inhibition of voltage-dependent outer hair cell motility and capacitance,'' J. Neurophysiol. **11**, 3096–3110.

Santos-Sacchi, J. (**1993**). ''Harmonics of outer hair cell motility,'' Biophys. J. **35**, 2217–2227.

Santos-Sacchi, J., and Dilger, J. P. (**1988**). ''Whole cell currents and mechanical responses of isolated outer hair cells,'' Hear. Res. **65**, 143–150.

Tolomeo, J. A., and Steele, C. R. (**1995**). ''Orthotropic piezoelectric properties of the cochlear outer hair cell wall,'' J. Acoust. Soc. Am. **97**, 3006–3011.

Zweig, G. (**1991**). ''Finding the impedance of the organ of corti,'' J. Acoust. Soc. Am. **89**, 1229–1254.

Effects of hearing impairment and presentation level on masking period patterns for Schroeder-phase harmonic complexes

Van Summers
Army Audiology & Speech Center, Walter Reed Army Medical Center, Washington, D.C. 20307-5001

(Received 22 October 1999; revised 20 January 2000; accepted 24 August 2000)

Masking period patterns (MPPs) for Schroeder-phase harmonic complexes containing equal-amplitude harmonics of a 100-Hz fundamental were determined for 5-ms tonal probes at 4000 and 1000 Hz. Maskers consisted of harmonics 2–50 (200–5000 Hz bandwidth) for 4000-Hz probes and harmonics 2–20 (200–2000 Hz) for 1000-Hz signals. Masked thresholds were determined for probe onsets 153, 155.5, 158, 160.5, and 163 ms following masker onset (masker duration=460 ms). Overall, results were similar for both probe frequencies. For listeners with normal hearing, MPPs for positive Schroeder-phase complexes masking 60 dB SPL probes were highly modulated and became flatter when probe level was increased to 80 dB SPL. MPPs were less modulated for listeners with sensorineural hearing loss than for normally hearing listeners at both 60 and 80 dB SPL probe levels. Thresholds in negative Schroeder-phase maskers were more similar across the two groups of listeners and across differences in probe position and probe level. The findings support an interpretation involving differences in the shape of the basilar-membrane waveform generated by each masker and possible influences of nonlinear cochlear processing on these internal responses. For normally hearing listeners, 60 dB SPL probes were most difficult to detect when temporally positioned so that probe frequency and masker instantaneous frequency were closely matched. For 80 dB SPL probes and for hearing-impaired listeners, probes presented at these same positions were often more easily detected than probes at other positions. The latter result appears to involve benefit associated with in-phase addition of the probe to a portion of the masker similar to the probe in both frequency and phase. This benefit was reduced or entirely eliminated when probe phase was altered so that this in-phase addition did not occur. [S0001-4966(00)05311-X]

PACS numbers: 43.66.Dc, 43.66.Ba [SPB]

I. INTRODUCTION

Active processes within the cochlea, sometimes referred to as the ''active mechanism,'' are critical to normal hearing in mammals. It is generally believed that these processes involve the introduction of additional mechanical force into the motion of the basilar membrane (BM) as a result of motile properties of the outer hair cells (OHCs). The added gain amplifies and sharpens BM response in regions most sensitive to the frequencies present in the input signal, increasing both the absolute sensitivity and frequency selectivity of cochlear processing (Davis, 1983). The gain is level-dependent (i.e., nonlinear), and diminishes as input levels increase.

The strength of these active processes is reduced with sensorineural hearing loss involving OHC damage. The resulting loss of nonlinear gain leads to more linear peripheral processing and is apparently a critical factor underlying many of the differences in hearing performance between normally hearing (NH) and hearing-impaired (HI) listeners (Moore, 1995). Several recent studies involving periodic stimuli known as ''Schroeder-phase'' harmonic complexes suggest that the effectiveness of these stimuli as maskers may be particularly sensitive to influences of the active mechanism (Carlyon and Datta, 1997a; Summers and Leek, 1998). The current experiments examined the masking properties of these harmonic complexes across test conditions where the influence of the active mechanism was expected to vary significantly.

Temporal waveforms for two pairs of Schroeder-phase complexes, used in the present study, appear in Fig. 1. Each ''positive'' Schroeder-phase masker (m_+) and the corresponding ''negative'' Schroeder-phase stimulus (m_-) are composed of the same harmonic components with all components set equal in amplitude [components 2–50 of a 100-Hz $F0$ in panels (a) and (b), components 2–20, 100-Hz $F0$ in panels (c) and (d)]. Phase assignments for components in the m_- complexes were determined from the following formula which is based on formulas given in Schroeder (1970),

$$\theta_n = -\pi n(n-1)/N,$$

where θ_n is the phase in radians of component n and N is the total number of harmonic components. The m_+ complex is produced by changing the minus sign at the beginning of the equation to a plus sign, resulting in a time-reversed version of the m_- waveform.

M_+ and m_- stimuli of equivalent bandwidth have identical long-term magnitude spectra and similar (very flat) temporal envelopes. However, the internal representations of these stimuli are apparently quite distinct. Modeling of the BM response to the stimuli in panels (c) and (d) of Fig. 1 (Smith *et al.*, 1986; Kohrausch and Sander, 1995) suggests

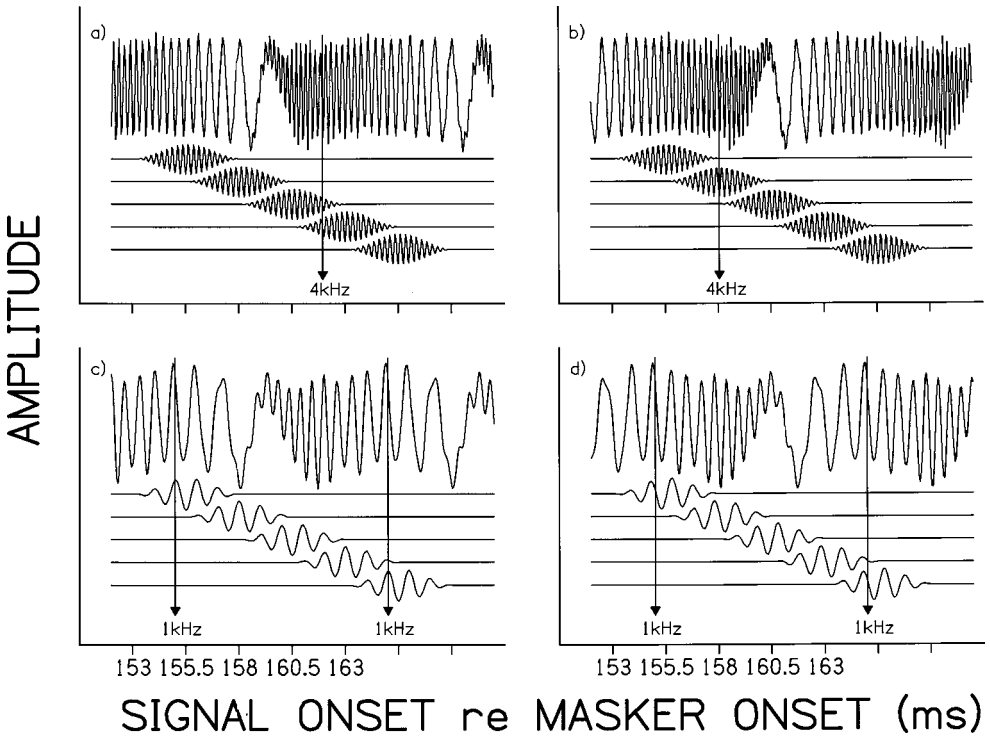


FIG. 1. Panels show brief portions of each Schroeder-phase complex and the five temporal alignments of the probe signals relative to that masker. M_+ complexes appear in the left-hand panels and m_- complexes are to the right. Stimuli used to test performance at 4000 Hz appear in the upper panels and 1000-Hz stimuli appear in the lower panels. Arrows in each panel indicate temporal locations where masker instantaneous frequency was equal to probe frequency.

that the m_+ complex produces a more modulated pattern of response at many locations on the BM, containing high-energy peak regions and low-energy valleys. The m_- complex produces a less modulated internal response. Recent physiological measurements on chinchillas are consistent with the modeling results, showing more modulated BM responses to m_+ than m_- stimuli (Recio and Rhode, in press).

The modulated BM response produced by the m_+ complex is apparently due to interactions between the phase structure of the complex and phase-dispersive properties of the membrane. Physiological data indicate that the impulse response of a fixed BM location shows a frequency-modulated glide in which low frequencies precede higher-frequency components of the response (de Boer and Nuttall, 1997; Recio *et al.*, 1998; de Boer and Nuttall, 2000a, 2000b). As seen in Fig. 1, m_+ complexes demonstrate the opposite sort of temporal structure: each pitch period contains a frequency glide in which high frequencies precede lower frequencies. To the extent that these opposing phase structures counteract one another, the components influencing a given point on the BM will become temporally aligned (i.e., equivalent in phase), resulting in a high-level peak within each cycle of BM movement.

Consistent with a modulated internal response to m_+ complexes, Kohlrausch and Sander (1995) reported that thresholds for brief 1100-Hz probes masked by the m_+ complex in Fig. 1(c) varied by approximately 17 dB depending on the temporal position of the probe within the 10-ms masker period. For the corresponding m_- complex, masked thresholds varied by only about 5.5 dB across positions within the masker period, i.e., the masking period pattern (MPP) was much flatter.

All of Kohlrausch and Sander's MPP data were collected at similar test levels (overall masker levels of 70 or 75 dB SPL, between approximately 51 and 66.5 dB/component). By replicating the Kohlrausch and Sander MPP experiment at a range of presentation levels, Carlyon and Datta (1997b) demonstrated that the modulated MPPs produced by m_+ maskers at moderate levels (masker level=69 dB/component) become much flatter at low levels (29 dB/component). Carlyon and Datta suggested that the greater modulation in the MPPs at their highest test levels may be due to broadening of auditory filters as levels increase. Assuming that component phases in the m_+ complex are appropriate to compensate for the BM phase response, broader filters will pass a larger number of components and should produce a peakier internal response. This could then account for more modulated MPPs at higher presentation levels. By this account, MPPs should become even more modulated at still higher presentation levels (based on further filter broadening), and MPP modulation should be greatest at the highest tested levels. Similarly, more modulated MPPs might be expected for listeners with sensorineural hearing loss than for NH listeners given that filtering is generally broader than normal for HI listeners (setting aside other possible influences of hearing loss on auditory processing). Note that if increases in MPP modulation are due to filter broadening, these increases are closely linked to active mechanism influences, given that filter broadening is generally assumed to be a consequence of reduced active processing (as a consequence of high test levels, hearing loss, or both).

Results reported by Summers and Leek (1998) using long-duration signals masked by Schroeder-phase complexes lead to different predictions about how hearing loss and high presentation level might influence MPPs for these maskers. In their results, the largest differences in the masking effectiveness between m_+ and m_- maskers were seen for NH listeners tested at moderate presentation levels: the two maskers were more similar in effectiveness at high presentation levels and in HI listeners. Summers and Leek interpreted

TABLE I. Ages and test ear audiometric thresholds for HI listeners (dB HL *re*: ANSI, 1989).

		Frequency (Hz)							
Subject	Age	250	500	1000	1500	2000	3000	4000	6000
PH	77	15	15	10	5	0	10	40	65
ED	74	10	10	10	10	5	40	30	35
JB	67	15	20	15	25	30	35	25	35
EF	61	25	30	35	40	40	40	30	30
DW	75	15	35	50	50	45	50	55	95
SK	69	40	30	25	25	30	55	55	75
CB	67	10	20	25	30	45	45	55	70
RL	77	40	50	55	65	60	55	55	70
AM	64	10	20	40	45	35	40	35	50

their results as indicating that large differences in the effectiveness of the two maskers require test conditions where the BM input–output function is nonlinear, i.e., NH listeners tested at moderate levels.

The present paper extends the work reported in Carlyon and Datta (1997b) and Summers and Leek (1998) by examining MPPs for m_+ and m_- maskers presented at moderate and high test levels to NH and HI listeners. If modulation in the m_+ MPP is due to broad auditory filters, the greatest modulation should occur for HI listeners and at high presentation levels. Conversely, if modulation in this MPP requires nonlinear cochlear processing, greater modulation is predicted for NH listeners tested at moderate levels than in other conditions.

II. EXPERIMENT 1

A. Methods

1. Subjects

Seven NH and nine HI listeners participated in experiment 1. Not all subjects were available for a sufficient number of sessions to allow testing of both the 4000- and 1000-Hz probes. For the 4000-Hz probes, seven NH and seven HI listeners were tested. For 1000-Hz probes, six NH and six HI listeners participated. [The specific HI listeners tested at each frequency are indicated in Fig. 5 (4000 Hz) and Fig. 8 (1000 Hz).]

NH listeners had thresholds $\leqslant$20 dB HL (*re*: ANSI, 1989) for audiometric frequencies between 250 and 4000 Hz. NH listeners ranged in age from 27 to 68 years with a mean age of 47.9.

Age and audiometric thresholds for the nine HI listeners appear in Table I. HI listeners had longstanding bilateral hearing losses. Six of the HI listeners had elevated thresholds of at least 25 dB HL for audiometric frequencies between 1000 and 6000 Hz. Three HI listeners had thresholds within normal limits at frequencies $\leqslant$1000 Hz with elevated thresholds at higher frequencies only. HI listeners ranged in age between 61 and 77 years with a mean age of 70.9. Bone conduction and/or immittance audiometry was consistent with cochlear site of lesion for all HI listeners.

2. Stimuli

Signals were 4000- and 1000-Hz tone pips with a total duration of 5 ms including 2.5-ms cosine^2 onset and offset ramps (no steady-state regions). Masking stimuli were m_+ and m_- complexes containing all harmonics from 200 to 5000 Hz (4000-Hz signal) or 200 to 2000 Hz (1000-Hz signal) of a 100-Hz fundamental, with harmonics set equal in amplitude. Maskers were 460 ms in duration including 2.5-ms cosine^2 onset and offset ramps. Signals were added to maskers in phase with the corresponding frequency component within the masker. Signal onset followed masker onset by 153, 155.5, 158, 160.5, or 163 ms, sampling performance at 2.5-ms intervals across one (10-ms) fundamental period within each masker.

Panels (a) and (b) of Fig. 1 show temporal waveforms for the m_+ and m_- complexes used to mask the 4000-Hz probes along with waveforms representing the probe signals at each temporal position tested. Panels (c) and (d) show masker and probe signals used to test performance at 1000 Hz. Previous MPP results suggest that the peak in the MPP (most effective masking) will occur when the probe is presented at a temporal location where the probe frequency and the instantaneous frequency of the masker coincide (Kohlrausch and Sander, 1995). Apparently this aligns the probe with the peak region of the internal response to the masker. The dashed lines in each panel indicate the point where the instantaneous frequency of each complex is equivalent to the probe frequency.

3. Procedure

Masker and signal stimuli were generated and combined digitally and converted to analog form by a D/A converter (TDT DD1) at a sample rate of 20 000 Hz. Waveforms were low-pass filtered at 8.6 kHz (TDT FT5, $>$70 dB/octave rolloff above 8.6 kHz), attenuated as required (TDT PA4), and led via a headphone buffer to a TDH-49 earphone in a sound-treated booth. Stimuli were presented to subjects monaurally.

Probe frequency was held constant and probe level was fixed at 60 or 80 dB SPL within a given block of trials. An adaptive three-interval forced-choice (3IFC) procedure was used to determine the overall masker level associated with 71% correct detection of each probe at each probe level (Levitt, 1971). A trial consisted of a random ordering of three stimulus intervals, with two containing the masker alone and one containing probe+masker. The interstimulus interval within a trial was 550 ms. Subjects indicated which interval contained the probe by touching the appropriate area on a touch-screen terminal. The terminal display provided correct-answer feedback following each response. On the first trial of each block, overall masker level was set 10 dB below the signal level. Masker level was then increased by 3 dB whenever two correct responses occurred consecutively and was decreased by 3 dB after each incorrect response. After five reversals in the adaptive track, the 3-dB step size was reduced to 1 dB and the track continued for an additional ten reversals. The threshold estimate for a trial block was the average of the masker levels at the final ten reversals

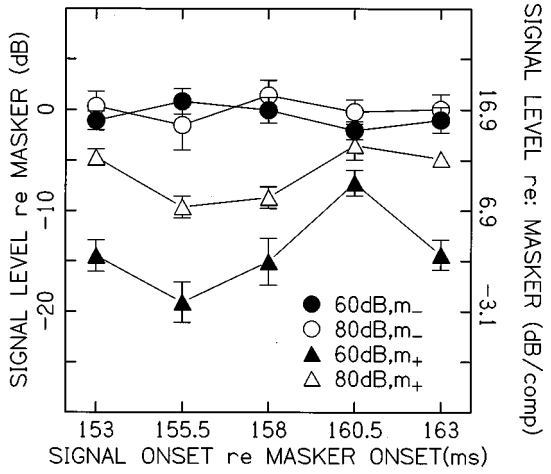


FIG. 2. Mean S/M levels at masked threshold for 4000-Hz probes: NH listeners ($N=7$). Bars indicate standard errors.

in the track. The data reported are averages from two trial blocks for each probe-frequency, probe-level condition. The standard deviation of these two threshold estimates was less than 6 dB in every instance, and less than 3 dB for 87% of the estimates.

Subjects only participated in conditions in which the probe signal was clearly audible in quiet. An in-quiet pretest was carried out for each probe frequency at each test level using the same 3IFC paradigm. Each subject was required to identify the probe interval correctly on ten out of ten trials to participate in a given test condition. In addition, testing of a given subject on a particular stimulus condition was discontinued if the masker reached 100 dB SPL and the probe remained audible to the subject. Masked thresholds were not measured in these instances.

B. Results

1. 4000-Hz probe

a. NH listeners. Figure 2 shows mean signal-to-masker (S/M) levels at masked threshold for the seven NH listeners tested on the 4000-Hz probe. The abscissa indicates delay in probe onset relative to masker onset. For all probe delays and at both probe levels, the m_- complex (circles) provided more masking than the m_+ masker (triangles). For the m_- masker, neither probe level nor position of the probe within the masker period had a clear influence on performance for NH listeners as a group: mean thresholds were near 0 dB S/M for all probe delays at both probe levels. However, effects of probe position on threshold were observed for some of the NH listeners at the 80-dB probe level. These individual results are discussed more fully below.

For the m_+ masker, probe level and probe position each had a clear effect on performance. For the 60-dB probe, mean thresholds differed by approximately 12 dB depending on the location of the probe within the period of the complex. For the 80-dB probe, thresholds showed a similar, but less modulated pattern, with a difference of about 6 dB across probe positions. At all probe locations, the m_+ masker was more effective at the 80-dB probe level than at the 60-dB level. This increase in masking effectiveness with level was greatest for positions where the probe was most easily detected (e.g., the 155.5-ms position). The location of the peak in the MPP for the m_+ complex (probe onset $=160.5$ ms) is consistent with greater masking when probe frequency and instantaneous frequency of the masker coincide [Fig. 1, panel (a)].

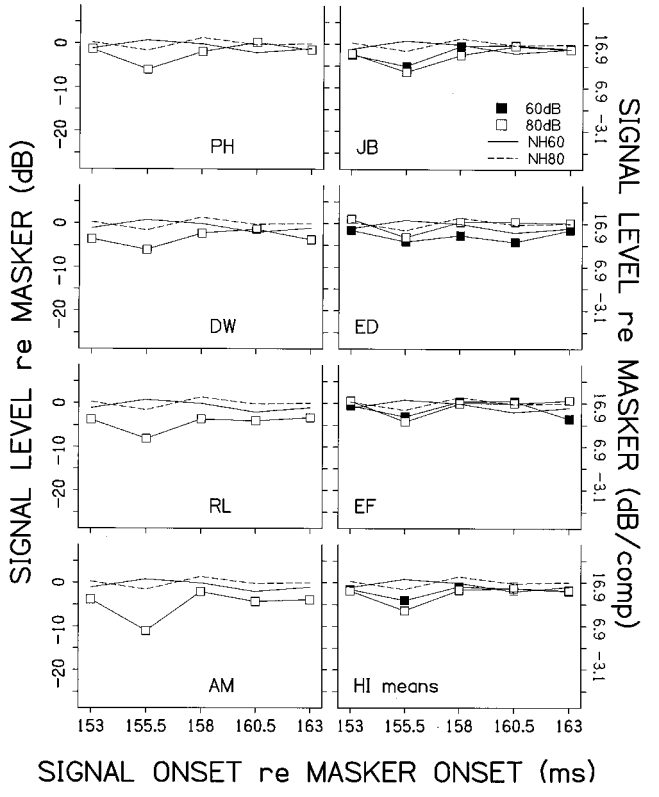


FIG. 3. Masked thresholds for 4000-Hz probes masked by m_- complexes: HI listeners. The lower right-hand panel shows means across listeners (bars indicate standard errors). Thresholds for individual HI listeners appear in the other panels. Mean HI thresholds for 80-dB probes are based on seven listeners, means for the 60-dB probes are based on three listeners (see text). Dashed and solid lines represent means for NH listeners on the m_- masker for 80- and 60-dB probes, respectively.

b. HI listeners. Seven HI listeners could reliably detect the 4000-Hz probe at 80 dB SPL in quiet and were tested at this probe level. Three of these listeners were able to detect the 4000-Hz probe reliably at 60 dB SPL in quiet and were also tested on this level.

Results for each HI listener and the mean results across listeners on the 4000-Hz probe are shown as squares in Fig. 3 (m_- masker) and Fig. 4 (m_+ masker). Seven of the panels in each figure show masked thresholds for a single HI listener along with mean data for the NH listeners using the same masker. The lower right-hand panels in each figure show mean results across HI listeners.

For HI listeners, the m_- complex was a less effective masker when the probe was positioned at the 155.5-ms location than at other placements. This result held for all seven listeners at the 80-dB probe level and was also seen in the mean data for the 60-dB probe (for the three HI listeners tested at this level). Although mean MPPs for NH listeners on the m_- masker were fairly flat, for the 80-dB probe, four of the NH listeners also showed lowest masked thresholds for the 155.5-ms probe position. For these four listeners the

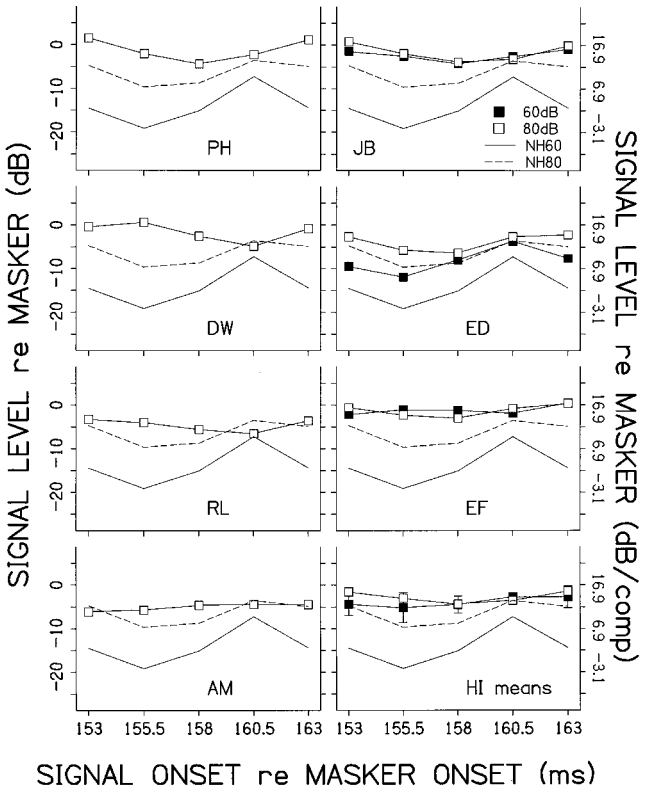


FIG. 4. Masked thresholds for 4000-Hz probes masked by m_+ complexes: HI listeners. Same format as Fig. 3.

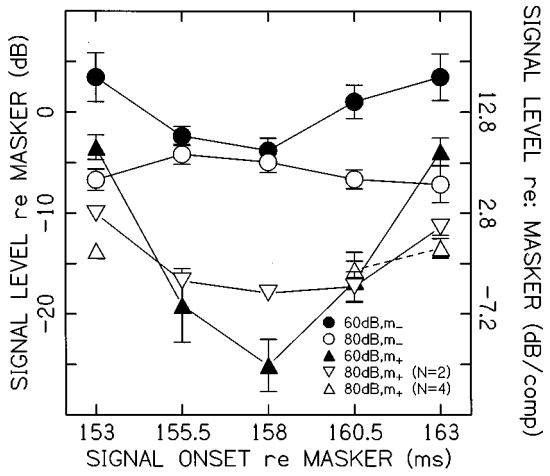


FIG. 5. Mean S/M levels at masked threshold for 1000-Hz probes: NH listeners. Thresholds for 60-dB probes and for m_- maskers using 80-dB probes are based on six NH listeners. For 80-dB probes masked by m_+ complexes, mean thresholds are shown separately for two NH listeners who provided data at all probe positions and for four NH listeners who only provided data at three positions (see text).

mean difference in threshold between the 155.5-ms probe position and the position producing the greatest masking was approximately 5 dB, similar to the value for HI listeners. However, in contrast to the pattern for HI listeners, this modulation in the MPPs for NH listeners was level dependent: lower thresholds for the 155.5-ms probe position were only observed for the 80-dB probe.

For the m_- masker, mean thresholds were similar for the NH and HI groups at four of the five probe locations. For the 155.5-ms location, where HI listeners showed their lowest thresholds, mean thresholds were actually lower by about 5 dB for the HI group than for the NH group (at both probe levels).

The instantaneous frequency of the m_- masker matched the 4000-Hz probe frequency near the temporal center of the probe for the 155.5-ms probe-onset condition [Fig. 1(b)]. Therefore, the lower masked thresholds at this probe position are not consistent with the expectation that the masker will be most effective when probe frequency and masker instantaneous frequency coincide. This reversal from the expected result was most evident in the MPPs for the HI listeners but also was observed for a number of NH listeners in testing at the 80-dB probe level. A possible basis for these reversals will be explored in experiment 2 (see Sec. III).

For the m_+ complex, MPPs for HI listeners were not as modulated as for NH listeners (Fig. 4). For the seven HI listeners tested on the 80-dB probe, the mean change in threshold across the MPP was about 3 dB, less than half the difference seen for NH listeners at the same probe level. Masked thresholds were generally higher for HI listeners than for NH listeners in the m_+ masker. In particular, for the 155.5-ms probe location, where thresholds for NH listeners were lowest, all NH listeners showed lower masked thresholds than any of the HI listeners. Based on thresholds for the three HI listeners tested on both 60- and 80-dB probe signals, probe level had less of an effect on thresholds in the m_+ masker for HI listeners than for NH listeners.

2. 1000-Hz probe

a. NH listeners. Six NH listeners were tested on 1000-Hz probes. Figure 5 shows mean S/M levels at masked threshold for these listeners on each masker. Assuming a fairly flat internal response to the m_- complex, it was expected that probe position and probe level would have little influence on masked thresholds for this masker. However, for the 60-dB probe, the MPP for this masker was modulated with mean thresholds varying by more than 7 dB depending on probe location. At the 60-dB probe level, masked thresholds were lowest for the 155.5- and 158-ms probe locations. Thresholds at these two locations showed little change with probe level. However, for the other probe positions, thresholds improved by approximately 8–11 dB when probe level was increased to 80 dB SPL. With these level-dependent changes in threshold, the MPP at 80 dB was fairly flat, showing only about 3 dB of variability in mean thresholds across probe positions.

For the m_+ complex, the temporal location of the 60-dB probe within the masker had a strong influence on performance for NH listeners. Mean thresholds varied by approximately 22 dB between probe delays producing the most (153- and 163-ms) and least (158-ms) effective masking. For four of the six NH subjects, the m_+ complex could not effectively mask the 80-dB probe at the 155.5- and 158-ms probe positions with exceeding the 100 dB SPL masker limit (S/M$=-20$ dB). In these two instances, masked thresholds could only be obtained for NH listeners SH and JC. Similar to the results for the m_- masker, the MPP for the m_+ complex was less modulated at the 80-dB probe level than at 60

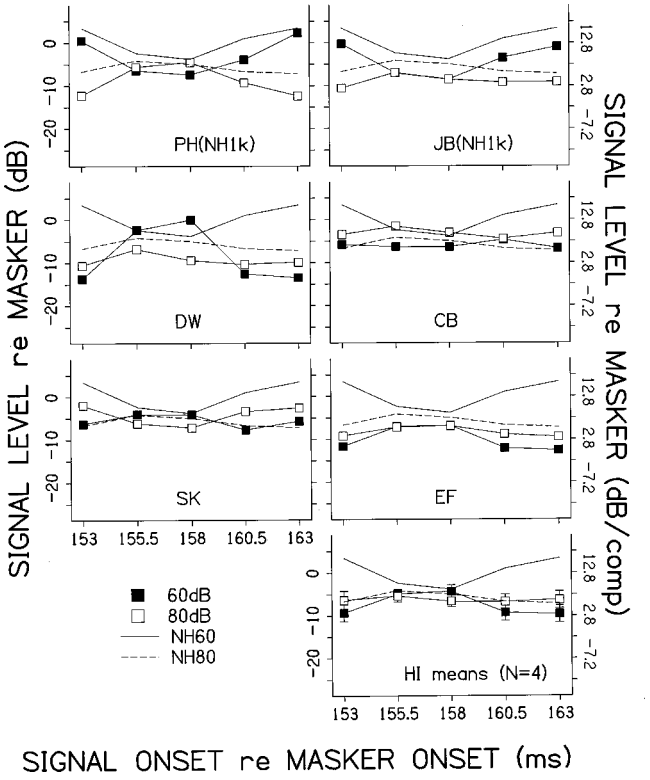


FIG. 6. Masked thresholds for 1000-Hz probes masked by m_- complexes: HI listeners. Same format as Fig. 3. The listeners whose data appear in the two upper panels had thresholds within the normal range at 1000 Hz. Mean values for HI listeners plotted in lower right-hand panel are based on the other four listeners.

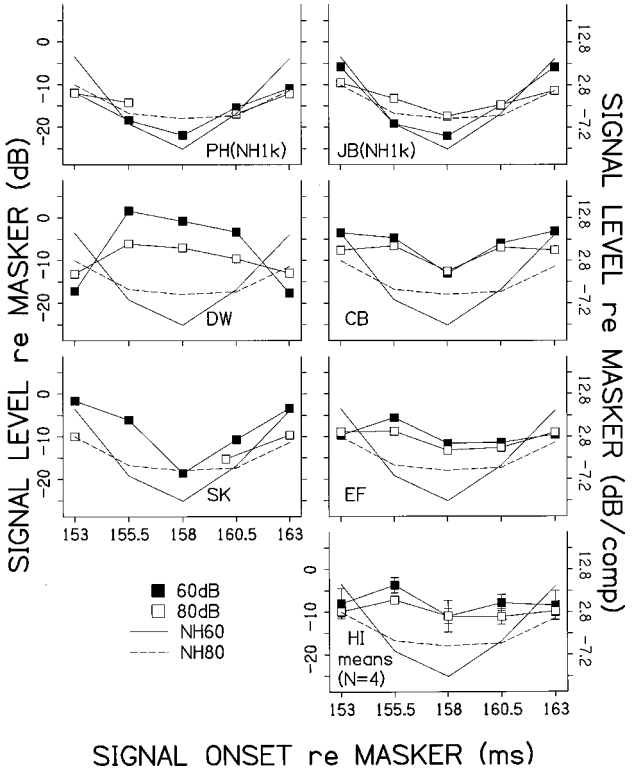


FIG. 7. Masked thresholds for 1000-Hz probes masked by m_+ complexes: HI listeners. Same format as Fig. 3. As in Fig. 6, mean values plotted in the lower right-hand panel do not include the results for PH and JB and are based on four listeners. Dashed and solid lines representing mean thresholds for NH listeners are based on six NH listeners for the 60-dB probes and two NH listeners for 80-dB probes (see text and Fig. 5).

dB, although this interpretation is complicated by the inability to test all of the NH listeners at all probe positions for the 80-dB probe. Overall, for NH listeners, the m_+ complex was a less effective masker than the m_- complex across probe positions at each presentation level.

Across probe levels, MPPs for the m_+ complexes showed highest masked thresholds for probe onsets of 153 and 163 ms. The same probe positions produced highest thresholds for the m_- complex in masking the 60-dB probes. These results are all consistent with more effective masking when probe frequency and masker instantaneous frequency coincide [Figs. 1(c), (d)]. A possible exception to this pattern is seen for the m_- masker and the 80-dB probes. The MPP for this condition, although quite flat, showed a tendency toward slightly *less* effective masking at the 153 and 163 probe onsets than at other probe positions. Note that this apparent reversal from the expected result is again at the 80-dB probe level, where a similar reversal was seen for some of the NH listeners in testing the 4000-Hz probes.

b. HI listeners. Six HI listeners who could reliably detect the 1000-Hz probe in quiet at 60 dB SPL were tested at both the 60- and 80-dB probe levels. Results appear in Figs. 6 (m_- complex) and 7 (m_+ complex).

Two of these listeners (PH and JB) had audiometric thresholds within normal limits ($\leqslant$20 dB HL) for frequencies up to 1000 Hz with elevated thresholds at higher frequencies. For both maskers and at both probe levels, the MPPs for these two listeners are similar and resemble the average patterns for NH listeners. In particular, note that in Fig. 6, MPPs for these two listeners show a reversal in shape with probe level similar to the pattern for the NH group: probes presented at delays of 153 and 163 ms go from being most readily masked for the 60-dB presentation level to most easily detected when presented at 80 dB.

Masked thresholds for the four HI listeners who had elevated thresholds at 1000 Hz did not resemble the results for the NH group. The mean MPPs for HI listeners shown in the bottom panels of Figs. 6 and 7 are based on the data for these four listeners. For the m_- complex and the 60-dB probes, the mean MPP for the HI listeners showed a different pattern of modulation than observed for the NH group. Although both groups showed MPPs with similar amounts of modulation, the probe positions producing the lowest masked thresholds for the NH group (155.5 and 158 ms) produced highest thresholds for the HI group and vice versa. For the 80-dB probe, the MPPs for HI listeners were fairly flat and the mean pattern approximated the results for the NH group. Probe level generally had less effect on performance for HI listeners than for NH listeners. For each group, the largest effects of probe level were seen for the 153-, 160.5, and 163-ms probe positions. At these locations, masked thresholds for the HI group worsened by between 2 and 4 dB when probe level was increased from 60 to 80 dB SPL. Conversely, for the NH group, thresholds improved by about 10 dB at these probe locations when level was increased. For these three probe positions, the HI listeners were able to detect the 60-dB probe at about 10–12 dB less favorable

S/M levels than the NH group. These group differences were not observed for the 80-dB probe with both groups showing similar, flat MPPs.

HI listeners showed a variety of MPPs for the m_+ masker (Fig. 7). The listeners with normal thresholds at 1000 Hz (PH, JB) and listener SK all showed patterns of masked thresholds which resembled the MPPs for the NH group. Listener DW showed reversals from the patterns observed for the NH listeners similar to the reversals already described for the m_- complex. The largest and most consistent difference between masked thresholds for the HI listeners and the mean NH data involves the 60-dB probe presented at the 158-ms probe delay, where the mean threshold for NH listeners was −25.1 dB. Masked thresholds for the HI listeners were all well above this value. Conversely, for probe positions where NH listeners showed their highest thresholds on the 60-dB probe (e.g., 153- and 163-ms positions), HI listeners outperformed the NH group: mean masked thresholds for the HI listeners at these probe locations were about 4.5 dB below the means for the NH group.

III. EXPERIMENT 2

The results of experiment 1 include instances where the probe position expected to provide most effective masking (based on a match between probe frequency and masker instantaneous frequency) actually provided the least masking. At these probe locations the probe is added to a portion of masker which is similar in both frequency and phase (note the similar relative phases of the probe and masker at the locations of the dashed lines in Fig. 1). Given that an in-phase addition of matching-frequency tones produces a greater increment in amplitude than the addition of tones differing in frequency (or, in the present context, adding the probe at some other position within the masker period), the in-phase addition of probe and masker may have contributed to lower masked thresholds when probe frequency and masker instantaneous frequency were aligned. This benefit would presumably be eliminated by altering the phase of the probe by 90 deg to prevent in-phase addition of signal +masker. This phase manipulation should not affect thresholds for probes aligned at other positions within the masker period where there is no frequency and phase correspondence between masker and probe. A brief follow-up experiment was carried out to test this hypothesis.

A. Methods

Two NH subjects and one HI subject tested in experiment 1 (JL, KG, and DW) participated along with one additional NH subject. The added subject (LK) had thresholds ≤20 dB HL (*re*: ANSI, 1989) for audiometric frequencies between 250 and 4000 Hz and was 33 years of age. The subjects previously tested in experiment 1 were retested here for probe conditions where a match between probe frequency and masker instantaneous frequency appeared to produce lower thresholds than at other probe positions in experiment 1 (i.e., the reversal from the expected pattern). The NH subjects were tested using 4000-Hz probes and the m_- masker (200–5000 Hz bandwidth). The HI subject was tested on

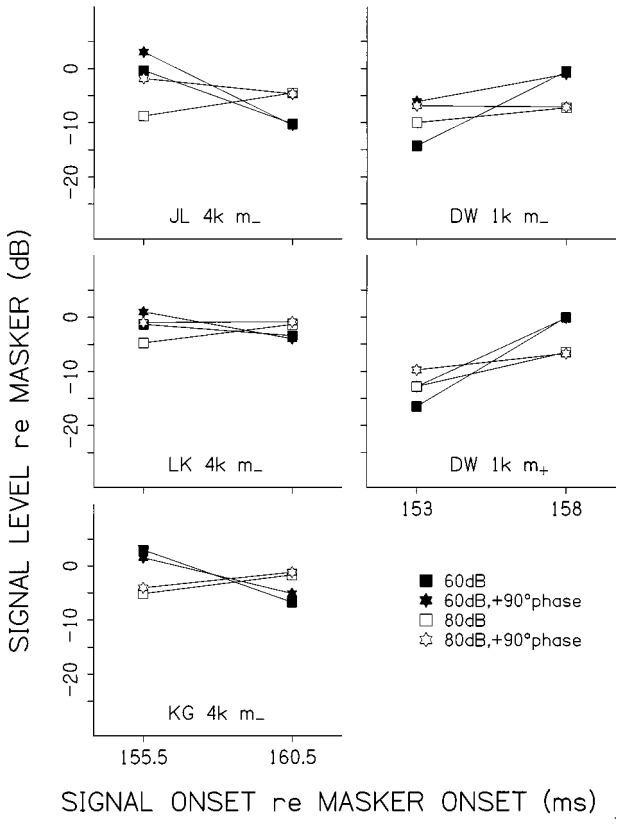


FIG. 8. Masked thresholds for probes using probe phases tested in experiment 1 and probes with phases advanced by 90 deg relative to original values. Left-hand column shows results for three NH listeners tested on 4000-Hz probes for the m_- complex. Right-hand column shows thresholds for HI listener DW tested on 1000-Hz probes for both m_- and m_+ maskers.

1000-Hz probes using both the m_- and m_+ maskers (200–2000 Hz bandwidth).

Masked thresholds were determined at two of the probe positions originally tested. These were: (1) the position where probe frequency and masker instantaneous frequency were closely aligned (the 155.5-ms probe-onset position for the test at 4000 Hz, the 153-ms position in the 1000-Hz data), and (2) positions 5 ms away from these locations, where probe frequency and masker instantaneous frequency did not coincide. Masked thresholds were determined for 60- and 80-dB probe levels with the probes used in experiment 1 and with new probes in which phases were advanced by 90 deg relative to their original values. Masked thresholds for each probe were determined using the identical procedures as in experiment 1.

B. Results

Masked thresholds for the three NH listeners appear in the left-hand panels of Fig. 8 and results for the HI listener appear in the right-hand portion of the figure. None of these data show an effect of probe phase on threshold when probe frequency and masker instantaneous frequency did not coincide (the 160.5- and 158-ms probe positions). However, for probe positions where probe frequency and masker instantaneous frequency overlap, probe phase did appear to influence performance in a number of instances. For listeners JL and

LK, thresholds for 80-dB probes presented at the 155.5-ms probe position were from 4–7 dB lower using the original probe phases (shown in Fig. 1) compared to performance with probe phases advanced by 90 deg. For these two subjects, masked thresholds for the 80-dB probes were 3–4 dB lower for the 155.5- vs 160.5-ms probe location using the original phase settings. This difference in threshold between the two locations was eliminated when probe phase was advanced by 90 deg. The effect of probe phase on threshold for these two listeners was reduced at the 60-dB probe level. The 4–7 dB difference seen for the 80-dB probes was about 2–3 dB for the 60-dB probes. As before, lower thresholds were seen using the original probe-phase settings than when probe phase was advanced.

The results for listener KG differed somewhat from those of the other NH listeners. For this listener, probe phase did not have a clear influence on performance for any of the tested conditions. Focusing on the results for the ''original-phase'' probes, all three NH listeners showed a level-dependent reversal in whether the 155.5- or the 160.5-ms probe position led to the higher thresholds. For the 60-dB probes, masked thresholds were 3–8 dB lower for the 160.5-ms position than at 155.5 ms. For the testing at 80 dB, thresholds were 3–4 dB higher at the 160.5-ms position than at the other location.

For the HI listener, probe phase also influenced masked thresholds for the probe location where probe frequency coincided with masker instantaneous frequency (the 153-ms position for both Schroeder-phase maskers). This effect was seen across probe levels and for both the m_+ and m_- maskers. In every instance the probe-phase settings used in experiment 1 led to lower masked thresholds than when probe phase was advanced by 90 deg. Differences in threshold between the 153- and 158-ms positions were consistently larger using the original probe-phase values than when probe phases were advanced by 90 deg. That is, similar to the results for JL and LK on 80-dB probes, it appeared that a substantial portion of the benefit observed when probe frequency matched masker instantaneous frequency over other probe alignments may have been due to ''in-phase'' addition of probe and masker given that this benefit was reduced when probe phase was altered.

IV. DISCUSSION

In experiment 1, for both 4000- and 1000-Hz signals, MPPs for NH listeners were more modulated for 60-dB probes masked by m_+ complexes than for any of the other combinations of probe level and masker. At both probe frequencies, the degree of MPP modulation for the m_+ complexes and the effect of probe level on MPP modulation for these maskers were greatly reduced for HI listeners compared to NH listeners. That is, MPP modulation was greatest in testing NH listeners at levels where cochlear processing is highly nonlinear. MPPs were much flatter for conditions linked to more linear processing (HI listeners, testing at levels $>$80–90 dB SPL). There are a number of ways in which nonlinear active processing in a healthy ear may have contributed to these results. However, prior to discussing these ideas, there is an alternative explanation of the data which should be addressed. For 80-dB probes masked by m_+ maskers, overall masker levels sometimes reached levels of 100 dB SPL. In some instances the probes remained audible even at this level. Therefore there is a possibility that the flattening of the MPPs at high levels for m_+ maskers was due to uncomfortably loud masker levels which lead listeners to respond incorrectly either deliberately or because the loudness became distracting. However, it is unlikely that this would explain the flattening of the MPPs for the 4000-Hz probes, given that mean thresholds remained below 90 dB SPL for all conditions (Fig. 2). For the 1000-Hz probes, masker loudness again does not appear to account for the overall pattern of flattening. At 1000 Hz, performance at three probe locations (153, 160.5, and 163 ms) was unchanged or actually tended to improve when probe level was increased to 80 dB SPL (Fig. 5). These results conflict with the prediction of poorer performance on the 80-dB probes due to uncomfortable loudness. Thus the results for 1000-Hz probes at the 155- and 158-ms probe locations appear to be the only data consistent with the prediction of poorer performance due to uncomfortable masker levels. It appears that flattening of the MPPs at high levels, seen for NH listeners at both probe frequencies, requires some other explanation.

In a study examining MPPs for similar stimuli, Carlyon and Datta (1997b) observed fairly flat MPPs for m_+ complexes presented at low test levels and more modulated MPPs at moderate levels. These authors successfully simulated their results using a transmission line model of cochlear processing (Giguère and Woodland, 1994). Carlyon and Datta suggested that greater modulation in the m_+ MPPs at higher levels could be the result of broader auditory filtering which allowed more masker components within the filter passband, and increased modulation in the internal response as a result [see Alcántara *et al.* (1996) for a similar conclusion based on modeling results for a related task]. This account would suggest that internal responses to the m_+ complexes (and resulting MPPs) should show the greatest modulation at the highest tested levels. However, the present study showed a reduction in MPP modulation for m_+ maskers at presentation levels higher than those tested by Carlyon and Datta (1997b).

The modeling approach described in Carlyon and Datta (1997b) was carried out using the 1000-Hz probes and the associated m_+ masker to determine whether this modeling would accurately simulate the flattening of m_+ MPPs for the 80-dB probe conditions (for NH listeners). Model parameters were set to simulate performance on 60- and 80-dB SPL probes for a healthy ear with normal OHC feedback, and in a passive system with OHC feedback turned off. Comparisons of the outputs from a single channel of the model (representing the response of a fixed BM location) in response to masker+probe and masker-alone inputs were used to simulate thresholds. The details of the modeling were essentially identical to those used previously (Carlyon and Datta, 1997b) and are described in the Appendix. Mean masked thresholds for NH listeners and simulated thresholds appear in Fig. 9. None of the simulated MPPs (unfilled symbols) showed as much modulation as the measured MPP for the 60-dB probes (filled triangles). The simulations showed in-

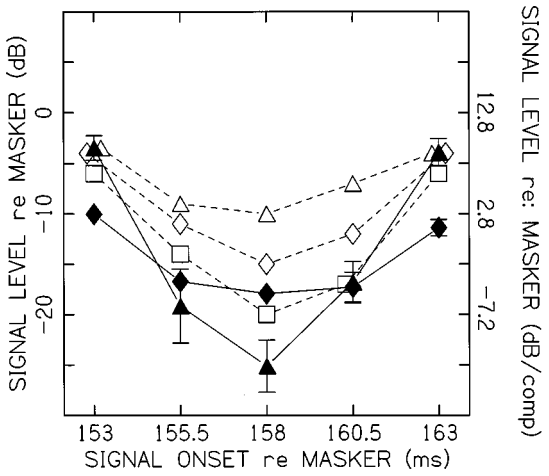


FIG. 9. Mean NH thresholds for 1000-Hz probes using m_+ maskers (▲=60 dB probes, ◆=80 dB probes) and simulated MPPs for the same stimuli (△=60 dB probes, normal OHC feedback; ◇=80 dB probes, normal OHC feedback; □=passive processing, no OHC feedback).

creased MPP modulation as probe level increased with the greatest modulation produced by the passive version of the model (unfilled squares). Thus, in the simulations, MPP modulation increased as filters broadened. This contrasts with the actual results where greater MPP modulation was seen at moderate levels for NH listeners than for conditions where auditory filtering was broader (higher test levels, HI listeners). The present behavioral results do not appear to be accurately simulated by the modeling approach described by Carlyon and Datta (1997b) and do not support the prediction of more modulated MPPs being directly linked to broader filtering.

The added masker components passed by broader filtering (at high levels and with hearing impairment) could actually decrease modulation in the BM response depending on the phase characteristics of these components. Measurements of BM response in both guinea pigs (Sellick *et al.*, 1983) and chinchillas (Ruggero *et al.*, 1997) show little or no phase dispersion in the propagation of frequencies more than half an octave below the characteristic frequency (CF) of a given BM location (constant slopes in phase-by-frequency plots, indicating pure delay independent of frequency). If low-frequency components in the present Schroeder-phase stimuli were processed in a similar way, these components could reduce modulation in the internal envelope (as they do in the external waveform). As filters broaden, the greater magnitude of these low-frequency components could decrease modulation in the BM response at a given place.

An alternative explanation of the present data is that probes presented in high-amplitude regions of a modulated internal response may be suppressed while probes placed in low-amplitude regions are not. By this account, the modulated MPPs for m_+ maskers would be linked to the patterns of ''phasic'' suppression which have been reported in physiological studies using low-frequency suppressor tones and higher-frequency probes (Rhode and Cooper, 1993; Cooper, 1996; Geisler and Nuttall 1997). In these studies, the BM response to a probe tone in the presence of a low-frequency masker (at a membrane location tuned to the probe frequency) shows a modulated pattern with much greater energy at the probe frequency during low-amplitude portions of the masker period (i.e., near zero crossings) than during high-amplitude portions. These results have been characterized as reflecting modulation of active gain based on displacement of the BM caused by the suppressor tone at a given moment (Cooper, 1996). That is, during low-amplitude portions of the suppressor, the active mechanism supplies frequency-selective gain which amplifies the representation of the probe. This gain is greatly reduced during high-amplitude portions of the suppressor and probe energy is reduced as a result. Psychoacoustic results which parallel these physiological data have also been reported. Zwicker (1976) reported changes in threshold of up to 16 dB for brief 2640-Hz probes placed at different positions within the period of a 100-Hz tonal masker. This threshold variability is qualitatively similar to the modulated MPPs reported here for m_+ maskers. In two-tone masking experiments similar to those reported by Zwicker, MPP modulation was greatly reduced in HI listeners relative to the masking patterns for NH listeners (Schroder, 1997). It seems likely that the variability in probe thresholds seen in these psychoacoustic studies (both the two-tone masking results and present data) is linked to the phasic suppression patterns observed in BM data, and therefore to cochlear nonlinearity [see Oxenham and Plack (1998) for related discussion].

Recently collected physiological data using Schroeder-phase complexes appear to support this possible link between MPP modulation in the present data and physiological suppression (Recio and Rhode, in press). In experiments examining BM responses in the chinchilla to m_+ and m_- waveforms, peakier response patterns were produced by m_+ than m_- complexes. Responses to near-CF probe tones showed phasic patterns of suppression in the presence of each complex with high-amplitude portions of the response envelopes associated with probe suppression. M_+ maskers produced suppression over less of the masker period than m_- stimuli. Suppression decreased as probe level increased, consistent with the reduced MPP modulation at high probe levels in the current data.

Changes in the BM phase response associated with changes in level and hearing status may have also contributed to the differences in performance across probe levels and between the two groups of listeners. If an altered BM phase response led to flatter internal responses to the m_+ maskers at high sound levels and in HI listeners, the observed flattening of the MPPs could be the result of these changes. The BM does show certain changes in phase response with input level (Ruggero *et al.*, 1997; Recio *et al.*, 1998). However, the overall response shows increasingly negative slope in the phase response with increasing frequency over levels ranging from 10 to 100 dB SPL and for both sensitive and insensitive cochleae (Nuttall and Dolan, 1996; Ruggero *et al.*, 1997). Thus the general pattern expected to produce a modulated internal response to the m_+ complex is present across levels and appears to be independent of cochlear status. In addition, physiological data indicate that effects of input level on BM phase response are difficult to demonstrate except in pristine cochleae (Ruggero

et al., 1997; Ruggero, personal communication) but that substantial nonlinearity may remain in somewhat less-sensitive preparations. This suggests that in the current experiments, level is likely to have had little influence on BM phase for even mildly impaired listeners. However, level-dependent MPP changes for listeners with mild impairments often tended to be similar in overall form to those of the NH listeners (ED in Fig. 4, SK and CB in Fig. 6), suggesting that level-dependent changes in BM phase may not completely account for the level effects reported in experiment 1.

Both psychoacoustic (von Klitzing and Kohlrausch, 1994; Baker *et al.*, 1998; Bacon *et al.*, 1999; Hicks and Bacon, 1999) and physiological data (Cooper and Yates, 1994) suggest that auditory processing may be more linear for low- than for high-frequency stimuli. If nonlinear processing played an important role in producing modulation in the current MPP data, it might be expected that greater modulation would be seen for the 4000- than the 1000-Hz probes. Instead, for the conditions expected to show the largest effects of nonlinear processing (i.e., NH listeners, 60-dB probes, and m_+ maskers), changes in masked threshold across the masker period were greater for 1000-Hz than for 4000-Hz probes. The prediction of more MPP modulation at 4000 Hz assumes that the m_+ complexes used to mask each probe were equally successful in producing a peaky BM response. The greater MPP modulation at 1000 Hz may indicate that the m_+ complex used to mask the 1000-Hz probes led to a peaker internal wave (i.e., more accurately counteracted BM phase response) than the m_+ complex used to mask the 4000-Hz signals. In addition, the transition to more linear auditory processing at low frequencies may occur below 1000 Hz in humans, so that the degree of nonlinearity between 1000 and 4000 Hz may be fairly similar (Bacon *et al.*, 1999; Plack and Oxenham, 2000).

Previous MPP results testing NH listeners and using similar Schroeder-phase complexes indicate that at moderate levels, these maskers are most effective when the probe is presented at a temporal location where probe frequency and masker instantaneous frequency coincide (Kohlrausch and Sander, 1995; Carlyon and Datta, 1997b). The present data for NH listeners tested at moderate levels are consistent with these earlier results. However, for HI listeners and for NH listeners tested at higher probe levels (i.e., conditions associated with fairly linear auditory processing), probes at this location were often more easily detected than at other positions. These lower masked thresholds appear to be related to an in-phase addition of the probe to a portion of the masker of similar phase and frequency. This effect may require test conditions allowing fairly linear auditory processing in that the result was never seen for NH listeners tested on the 60-dB probes. That is, the benefits provided by nonlinear active gain for probes presented in low-amplitude regions of a modulated internal response may outweigh the benefits of in-phase addition of the probe and masker for conditions where cochlear processing is nonlinear.

The current results showed large differences in the effectiveness of m_+ and m_- maskers under test conditions associated with nonlinear cochlear processing. The two maskers were more nearly equal in effectiveness for conditions associated with more linear processing (high levels, HI listeners). A number of factors linked to the cochlear amplifier, including possible suppressive effects and level-dependent changes in the phase and magnitude response of effective filtering, may have contributed to these differences. A more complete characterization of the contribution of these and other factors to the observed threshold differences will almost certainly require further physiological work using these types of stimuli along with more detailed modeling efforts.

ACKNOWLEDGMENTS

Thanks to Armin Kohlrausch, Sid Bacon, and an anonymous reviewer for valuable comments on a previous version of this article. Thanks to Ken Grant, Marjorie Leek, and Anna Schroder for valuable discussions of the results and to Franz Seitz for experimental programming. This research was supported by Grant No. DC 03553 from the National Institute on Deafness and Other Communication Disorders to Walter Reed Army Medical Center and by the Clinical Investigation Service, Walter Reed Army Medical Center, under Work Unit #2577. All subjects participating in this research provided written informed consent prior to beginning the study. The opinions or assertions contained herein are the private views of the author and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

APPENDIX: IMPLEMENTATION DETAILS OF MPP SIMULATION

Masked thresholds were simulated using the transmission-line model (TLM) of cochlear processing described by Giguère and Woodland (1994). Masker-alone and masker+probe waveforms were input to the TLM and thresholds were simulated based on a comparison of the outputs of a single channel (representing the response of a fixed location on the BM) to these two inputs. TLM outputs were rectified, low-pass filtered (6-ms Hanning window, equivalent rectangular duration=3 ms), and then compared based on division of the masker+probe output by the masker-alone output. The largest value in the resulting ratio waveform (MAXRAT) was taken as a measure of the change in response due to the presence of the probe signal. To begin, a masker+probe input was prepared representing the 1-kHz 60-dB probe at 153-ms delay *re*: the m_+ masker with S/M $=-3.46$ dB (the mean threshold for NH listeners on this stimulus). MAXRAT values comparing TLM responses to this input to responses for a masker-alone m_+ input were examined for 11 channels of the model (center frequencies from 751 to 1948 Hz in steps of 10%). The largest MAXRAT value occurred in the 1100-Hz channel and the outputs of this channel were used in the rest of the simulation. The MAXRAT value for this pair of inputs for the 1100-Hz channel was treated as the value required for the signal to be above threshold (THRESHRAT) under all masking conditions. Masker-alone and masker+probe stimuli were then prepared for S/Ms ranging from -1 to -22 dB for each of the five probe-delay conditions. For each probe delay, MAX-

RAT was determined beginning with the highest S/M and continuing with S/M, decreasing by 1 dB until MAXRAT was less than THRESHRAT. The simulation was discontinued at this point and the S/M for this final masker+probe input was used as the predicted threshold. To simulate normal hearing, the outer hair cell feedback parameter was set to its default value (feedback_tlf=0.99) and 60- and 80-dB probe conditions were simulated by appropriate manipulation of the dBwave parameter of the model, which allows specification of input wave level. In addition, output for a linear version of the model was determined by setting OHC feedback gain to zero (feedback_tlf=0.0). With OHC feedback turned off, the output of the model is level-independent and is therefore only examined at one presentation level.

Alcántara, J. I., Holube, I., and Moore, B. C. J. (**1996**). ''Effects of phase and level on vowel identification: Data and predictions based on a nonlinear basilar-membrane model,'' J. Acoust. Soc. Am. **100**, 2382–2392.

ANSI (**1989**). ''Specifications for audiometers,'' American National Standards Institute, New York.

Bacon, S. P., Boden, L. N., Lee, J., and Repovsch, J. L. (**1999**). ''Growth of simultaneous masking for $f_m<f_s$: Effects of overall frequency and level,'' J. Acoust. Soc. Am. **106**, 341–350.

Baker, R. J., Rosen, S., and Darling, A. M. (**1998**). ''An efficient characterization of human auditory filtering across level and frequency that is also physiologically reasonable,'' in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London).

Carlyon, R. P., and Datta, A. J. (**1997a**). ''Excitation produced by Schroeder-phase complexes: Evidence for fast-acting compression in the auditory system,'' J. Acoust. Soc. Am. **101**, 3636–3647.

Carlyon, R. P., and Datta, A. J. (**1997b**). ''Masking period patterns of Schroeder-phase complexes: Effects of level, number of components, and phase of flanking components,'' J. Acoust. Soc. Am. **101**, 3648–3657.

Cooper, N. P. (**1996**). ''Two-tone suppression in cochlear mechanics,'' J. Acoust. Soc. Am. **99**, 3087–3098.

Cooper, N. P., and Yates, G. K. (**1994**). ''Nonlinear input-output functions derived from the responses of guinea-pig cochlear nerve fibers: Variations with characteristic frequency,'' Hear. Res. **78**, 221–234.

Davis, H. (**1983**). ''An active process in cochlear mechanics,'' Hear. Res. **9**, 79–90.

de Boer, E., and Nuttall, A. L. (**1997**). ''The mechanical waveform of the basilar membrane. I. Frequency modulations ('glides') in impulse responses and cross correlation functions,'' J. Acoust. Soc. Am. **101**, 3583–3592.

de Boer, E., and Nuttall, A. L. (**2000a**). ''The mechanical waveform of the basilar membrane. II. From data to models—and back,'' J. Acoust. Soc. Am. **107**, 1487–1496.

de Boer, E., and Nuttall, A. L. (**2000b**). ''The mechanical waveform of the basilar membrane. III. Intensity effects,'' J. Acoust. Soc. Am. **107**, 1497–1507.

Geisler, C. D., and Nuttall, A. L. (**1997**). ''Two-tone 'low-side' suppression in responses of basal basilar membrane,'' in *Diversity in Auditory Mechanics*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, and E. Hecht-Poinar (World Scientific, Singapore).

Giguère, C., and Woodland, P. C. (**1994**). ''A computational model of the auditory periphery for speech and hearing,'' J. Acoust. Soc. Am. **95**, 331–342.

Hicks, M. L., and Bacon, S. P. (**1999**). ''Psychophysical measures of auditory nonlinearities as a function of frequency in individuals with normal hearing,'' J. Acoust. Soc. Am. **105**, 326–338.

Kohlrausch, A., and Sander, A. (**1995**). ''Phase effects in masking related to dispersion in the inner ear. II. Masking period patterns of short targets,'' J. Acoust. Soc. Am. **97**, 1817–1829.

Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Moore, B. C. J. (**1995**). *Perceptual Consequences of Cochlear Damage* (Oxford University Press, Oxford).

Nuttall, A. L., and Dolan, D. F. (**1996**). ''Steady-state sinusoidal velocity responses of the basilar membrane in the guinea pig,'' J. Acoust. Soc. Am. **99**, 1556–1565.

Oxenham, A. J., and Plack, C. J. (**1998**). ''Suppression and the upward spread of masking,'' J. Acoust. Soc. Am. **104**, 3500–3510.

Plack, C. J., and Oxenham, A. J. (**2000**). ''Basilar membrane nonlinearity estimated by pulsation threshold,'' J. Acoust. Soc. Am. **107**, 501–507.

Recio, A., and Rhode, W. S. (**2000**). ''Basilar membrane responses to broadband stimuli,'' J. Acoust. Soc. Am. **108**, 2281–2298.

Recio, A., Rich, N. C., Narayan, S. S., and Ruggero, M. A. (**1998**). ''Basilar-membrane responses to clicks at the base of the chinchilla cochlea,'' J. Acoust. Soc. Am. **103**, 1972–1989.

Rhode, W. S., and Cooper, N. P. (**1993**). ''Two-tone suppression and distortion production on the basilar membrane in the hook region of cat and guinea pig cochleae,'' Hear. Res. **66**, 31–45.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (**1997**). ''Basilar-membrane responses to tones at the base of the chinchilla cochlea,'' J. Acoust. Soc. Am. **101**, 2151–2163.

Ruggero, M. A. (personal communication).

Schroder, A. C. (**1997**). ''Masking period patterns for 40-Hz biasing tones in normal-hearing and hearing-impaired listeners,'' Ph.D. thesis, University of Minnesota.

Schroeder, M. R. (**1970**). ''Synthesis of low peak-factor signals and binary sequences with low autocorrelation,'' IEEE Trans. Inf. Theory **16**, 85–89.

Sellick, P. M., Yates, G. K., and Patuzzi, R. (**1983**). ''The influence of Mossbauer source size and position on phase and amplitude measurements of the guinea pig basilar membrance,'' Hear. Res. **10**, 101–108.

Smith, B. K., Sieben, U. K., Kohlrausch, A., and Schroeder, M. R. (**1986**). ''Phase effects in masking related to dispersion in the inner ear,'' J. Acoust. Soc. Am. **80**, 1631–1637.

Summers, V., and Leek, M. R. (**1998**). ''Masking of tones and speech by Schroeder-phase harmonic complexes in normally hearing and hearing-impaired listeners,'' Hear. Res. **118**, 139–150.

von Klitzing, R., and Kohlrausch, A. (**1994**). ''Effect of masker level on overshoot in running- and frozen-noise maskers,'' J. Acoust. Soc. Am. **95**, 2192–2201.

Zwicker, E. (**1976**). ''Psychoacoustic equivalent of period histograms [in memoriam Dr. Russell Pfeiffer],'' J. Acoust. Soc. Am. **59**, 166–185.

Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise

Brian R. Glasberg[a] and Brian C. J. Moore
Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, England

(Received 28 June 1999; revised 21 March 2000; accepted 7 August 2000)

Thresholds for detecting sinusoidal signals were measured as a function of the spectral width of a notch in a noise masker. The notch was positioned both symmetrically and asymmetrically around the signal frequency. The noise was designed to create equal excitation per ERB within its passbands (uniformly exciting noise), after allowing for the transfer function of the headphone and the middle ear. For a signal frequency of 250 Hz, the level per ERB ranged from 35 to 80 dB in 15-dB steps. For signal frequencies of 500, 1000, 2000, and 4000 Hz, the level per ERB ranged from 40 to 70 dB per ERB in 15-dB steps. Auditory filter shapes were derived from the data by modeling the auditory filter as the sum of a sharply tuned tip filter and a broader tail filter. The gain of the tip filter was assumed to be a function of level. The shape of the tip filter and the gain and shape of the tail filter were assumed to be level independent. The data for all levels were fitted simultaneously. The data were fitted best when the gain of the tip filter was assumed to be a function of the signal level (as opposed to the masker level per ERB). The filter shapes showed a level dependence that qualitatively resembled the level dependence of filtering on the basilar membrane. The maximum gain of the tip filter tended to increase with increasing center frequency up to 1 kHz, but to remain roughly constant for higher frequencies. © *2000 Acoustical Society of America.* [S0001-4966(00)03211-2]

PACS numbers: 43.66.Dc, 43.66.Ba [SPB]

I. INTRODUCTION

It is widely believed that the frequency selectivity measured behaviorally in masking experiments is largely determined by the filtering that occurs on the basilar membrane (BM) (Fletcher, 1940; Moore, 1986; Evans *et al.*, 1989). The tuning of a specific point of the BM is determined by two components: a passive component reflecting the hydromechanical properties of the BM and surrounding fluids and structures, and an active component reflecting the influence of the outer hair cells on BM motion. The passive component appears to be roughly linear and does not vary with level, while the active component is highly nonlinear (Rhode and Robles, 1974; Sellick *et al.*, 1982; Ruggero *et al.*, 1997). At low sound levels, the active component has a strong influence, and sharp tuning is observed, at least around the ''tip'' of the tuning curve; the relatively shallow low-frequency ''tail'' is thought to reflect the passive component. As the level is increased, the gain of the active tip filter decreases and the passive filter plays a greater role. Thus, the tuning becomes broader with increasing level, especially for frequencies below the characteristic frequency (CF) of the place on the BM whose response is being measured.

Auditory filter shapes estimated behaviorally using the notched-noise method have also been characterized as having a sharply tuned ''tip'' and a broader ''tail'' (Patterson *et al.*, 1982; Glasberg *et al.*, 1984; Rosen *et al.*, 1992, 1998; Baker *et al.*, 1998). Patterson *et al.* (1982) proposed that each side of the auditory filter can be described as the sum of two rounded-exponential functions, one characterizing the tip and one the tail. Usually, auditory filter shapes have been measured using a fixed noise level. When several fixed noise levels are used, the data for each noise level being analyzed separately, the derived filter tends to become broader on the low-frequency side as the level increases (Lutfi and Patterson, 1984; Patterson and Moore, 1986; Moore and Glasberg, 1987; Glasberg and Moore, 1990; Rosen *et al.*, 1992, 1998); changes on the high-frequency side are smaller, and are not consistent across studies.

Rosen and his co-workers (Rosen *et al.*, 1992, 1998; Baker *et al.*, 1998) have proposed a method for deriving auditory filter shapes from notched-noise data in which several different levels are used within the same experiment. The analysis of the results was based on the assumption that each side of the filter could be described as the sum of two rounded-exponential functions, but the parameters describing the filter shape were made functions of level. They found that the data were fitted better when the parameters were made to depend on the signal level than when they were made to depend on the masker spectrum level. The broadening of the low-frequency side of the filter with increasing level could be well described by a change in the gain of the tip filter relative to that of the tail filter.

Glasberg *et al.* (1999) have also described an approach to analyzing notched-noise data with a two-filter model. In their approach, which is described in more detail later in this article, the gain of the tip filter was assumed to vary with level in the same way as the BM gain, for a tone close to the CF; the gain was greater at low levels than at high levels (Robles *et al.*, 1986; Ruggero *et al.*, 1997). The relevant

[a]Electronic mail: bg12@cus.cam.ac.uk

level was assumed to be the input level per ERB, where ERB stands for the equivalent rectangular bandwidth of the auditory filter (Glasberg and Moore, 1990); the assumption that auditory filtering depends on the input level per critical bandwidth (essentially equivalent to the ERB) has also been made by Zwicker and his co-workers, in the calculation of excitation patterns (Zwicker, 1970; Zwicker and Fastl, 1990). Glasberg *et al.* used this model to analyze the data of Lutfi and Patterson (1984) and Rosen *et al.* (1998). The filter shapes derived from the notched-noise masking data showed a level dependence that qualitatively resembled the level dependence of filtering on the BM.

In this article, we present data obtained using a modified version of the notched-noise method. Previously, this method has often been used with noise having a uniform spectrum level within its passbands. We used a noise designed to produce equal excitation at each frequency within the passband of the noise. Such noise has sometimes been referred to as ''uniformly exciting noise (UEN)'' (Zwicker and Fastl, 1990). The noise was designed to have a constant level in each ERB after allowing for the frequency response of the earphone at the eardrum and the transmission of sound through the middle ear.

The use of UEN has two advantages compared to white noise. First, if it is assumed that the filter shape is controlled by the input level per ERB, then the analysis of the results to derive auditory filter shapes is greatly simplified. Second, with UEN, the output of the auditory filter is influenced by noise components both above and below the center frequency, even when one of the noise bands falls at a very low or high frequency, where it might be significantly attenuated by the middle ear. This helps to limit the degree of off-frequency listening (Patterson and Nimmo-Smith, 1980). The results are analyzed both using the method described by Glasberg *et al.* (1999) and using a modified version of the ''PolyFit'' method described by Rosen and co-workers (Rosen *et al.*, 1992, 1998; Baker *et al.*, 1998).

II. EXPERIMENT 1: DETECTION OF TONES IN UNIFORMLY EXCITING NOTCHED NOISE

A. Stimuli

Data were collected for signal frequencies, f_s, of 250, 500, 1000, 2000, and 4000 Hz. The data for the four highest signal frequencies were collected before the data for 250 Hz, and the calculation of the noise spectrum was based on slightly different assumptions for these two cases. In all cases, the goal was that the spectrum of the noise reaching the cochlea would have approximately a constant level per ERB within its passbands. We started by calculating the spectral shaping that would be applied to a white noise to give a constant level per ERB in the electrical signal. At a given frequency, to achieve a level per ERB of L_{ERB}, the spectrum level was set to $L_{\mathrm{ERB}} - 10\log_{10}(\mathrm{ERB})$; ERB values were taken from Glasberg and Moore (1990).[1]

Then a ''correction'' was applied for the transmission of sound through the headphone and the outer and middle ear. For signal frequencies of 500, 1000, 2000, and 4000 Hz, we assumed, following Glasberg and Moore (1990), that the overall effect of the transfer function from the free field to the cochlea was reflected in the shape of the 100-phon equal-loudness contour; essentially the transfer function was obtained by inverting the equal loudness contour. Since the Sennheiser HD414 earphones used had a ''free-field response'' (i.e., the response at the eardrum was similar to that produced by a flat-response loudspeaker in a free field), the transfer through the earphone to the cochlea could also be represented by the equal-loudness contour. The correction applied to the noise spectrum was the inverse of the assumed transfer function. For example, the transfer function had a peak around 2700 Hz, so the corrected noise had a dip around 2700 Hz.

For $f_s = 250$ Hz, a Sennheiser HD580 earphone was used, as this has a more extended low-frequency response and lower distortion than the HD414. The frequency response of the HD580 at the eardrum was estimated using a KEMAR manikin (Burkhard and Sachs, 1975), averaging the results for the large and small ears. The transfer function through the middle ear was assumed to have the form specified in Moore *et al.* (1997); this form was largely based on the physiological measurements of Puria *et al.* (1997). The inverse of the combined effect of the earphone and the middle-ear transfer functions was applied to the (electrical) noise spectrum.

Noises were created digitally, on a Silicon Graphics Indy computer, using a 25-kHz sampling rate and 16-bit precision. Each noise actually consisted of two passbands, each with a width of $0.4f_s$ (except when the low-frequency edge of the lower noise band would fall below 40 Hz, when the edge frequency was set to 40 Hz), with a spectral notch between the two passbands. The spectral shaping within each passband was determined as described above. The spectral magnitudes were specified at 0.1-Hz intervals. Component phases were random. A ten-second noise sample was then calculated using an inverse Fourier transform.

The notch width for each noise was defined as the deviation of each edge of the notch from f_s, divided by f_s, denoted $\Delta f_s/f_s$. There were seven conditions where the notch was symmetrically placed around f_s; values of $\Delta f_s/f_s$ were 0, 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6. In six conditions, $\Delta f_s/f_s$ was 0.2 units greater for the upper than for the lower band; values of $\Delta f_s/f_s$ for the lower band were 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6. Finally, in six conditions, $\Delta f_s/f_s$ was 0.2 units greater for the lower than for the upper band; values of $\Delta f_s/f_s$ for the upper band were 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6. For $f_s = 500$, 1000, 2000, and 4000 Hz, three levels per ERB were used: 40, 55, and 70 dB (effective level reaching the cochlea, assuming that the middle ear has an effective transmission of 0 dB at 1000 Hz). The corresponding spectrum levels at the eardrum for a frequency of 1000 Hz were 19, 34, and 49 dB. For $f_s = 250$ Hz, four levels per ERB were used: 35, 50, 65, and 80 dB.

For use during the experiment, the digital noise files were transferred to a PC. During a trial, three noise bursts were presented, each with 190-ms steady-state portions and 10-ms raised-cosine ramps. The interstimulus interval was 500 ms. The noise bursts were selected randomly from the 10-s noise buffer, and the ramps were applied digitally be-

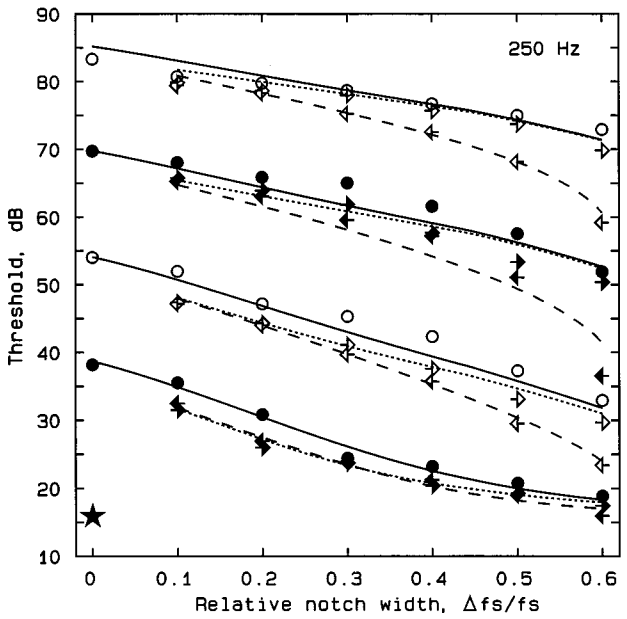


FIG. 1. Mean masked thresholds in dB SPL obtained for a signal frequency of 250 Hz. Thresholds are plotted as the effective level reaching the cochlea, allowing for the response of the headphone and the middle ear. The middle ear relative response was based on the function published in Moore *et al.* (1997), assuming a response of 0 dB at 1000 Hz. Thresholds are plotted as a function of relative notch width, $\Delta f_s/f_s$. The curves show thresholds predicted by the model described in the text. Circles (and solid lines) denote conditions where the notch was symmetrically placed about the signal frequency. Left-pointing arrows (and long-dashed lines) denote conditions where $\Delta f_s/f_s$ for the lower noise band was 0.2 greater than $\Delta f_s/f_s$ for the upper band. Right-pointing arrows (and short-dashed lines) denote conditions where $\Delta f_s/f_s$ for the upper noise band was 0.2 greater than $\Delta f_s/f_s$ for the lower band. The four groups of curves correspond to four masker levels per ERB: 35, 50, 65, and 80 dB (from bottom to top). The solid star indicates the absolute threshold for detecting the signal.

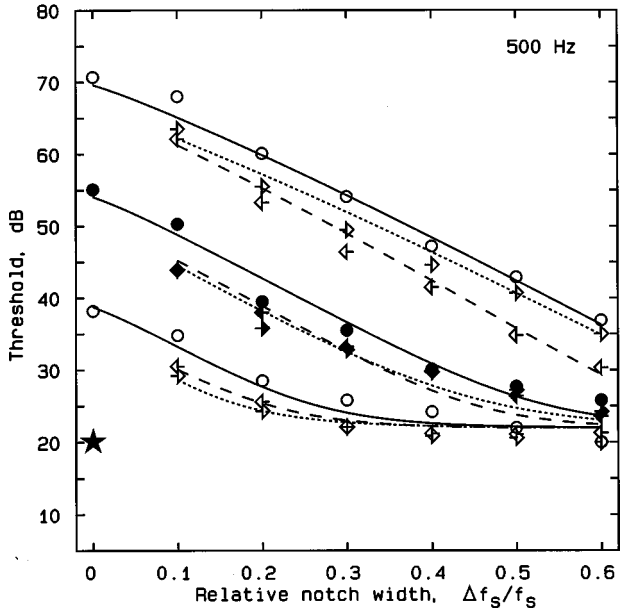


FIG. 2. As in Fig. 1, but for a signal frequency of 500 Hz. The levels per ERB were 40, 55 and 70 dB.

fore presentation using a Tucker-Davis Technologies (TDT) system with a 16-bit digital-to-analog converter (DD1). The sinusoidal signal was also generated digitally (25-kHz sampling rate), and it was presented synchronously with one of the three noise bursts, selected at random; the signal duration equaled the masker duration. The masker and signal were low-pass filtered separately at 10 kHz, using two sections of a Kemo VBF8 Mk 4 filter (90 dB/oct), attenuated (TDT PA4), summed (TDT SM3), passed through a headphone buffer (TDT HB6), and delivered to a double-walled sound attenuating booth. Stimuli were delivered to the headphone via a final manual attenuator (Hatfield 2125).

B. Method

Thresholds were measured using a three-alternative forced-choice three-down one-up procedure tracking the 79.4% point on the psychometric function. Observation intervals were marked by lights on the response box and feedback was provided after each trial by a light indicating the correct interval. Twelve turnpoints were obtained in a given run, and the threshold estimate for that run was taken as the mean value of the levels at the last eight turnpoints. The step size was 5 dB up to the first four turnpoints, and 2 dB thereafter. At least two runs were obtained for each condition. When the thresholds for those two runs differed by more than 2 dB, at least one additional run was obtained, and the results for all runs were averaged.

C. Subjects

Four subjects were tested using signal frequencies of 500, 1000, 2000, and 4000 Hz. One was author BG. BG and three different subjects were tested at 250 Hz. All subjects had absolute thresholds better than 10 dB HL at the standard audiometric frequencies and all had previous experience in psychoacoustic tasks. Their ages ranged from 24 to 54 years. Subjects were each given about 2 hours of practice, after which their performance appeared to be stable.

D. Results

The pattern of results was similar across subjects and the mean results for each signal frequency are given in Figs. 1–5. In each figure, the abscissa shows the smaller of the two values of $\Delta f_s/f_s$. Circles denote conditions where the notch was symmetrically placed about the signal frequency. Left-pointing arrows denote conditions where $\Delta f_s/f_s$ for the lower noise band was 0.2 greater than $\Delta f_s/f_s$ for the upper band. Right-pointing arrows denote conditions where $\Delta f_s/f_s$ for the upper noise band was 0.2 greater than $\Delta f_s/f_s$ for the lower band. The curves show predictions of a model that will be described later. Thresholds are specified as the effective level reaching the cochlea after allowing for the transfer through the headphone and the middle ear; thus, the signal levels are directly comparable to the levels per ERB of the noise maskers. The absolute threshold for detecting the signal, indicated by an asterisk, is expressed in the same way.

For $f_s = 250$ Hz (Fig. 1), for the lowest noise level (filled symbols at bottom) the left-pointing arrows lie very close to the right-pointing arrows, indicating that the auditory filter is almost symmetric. However, for the higher noise levels, the

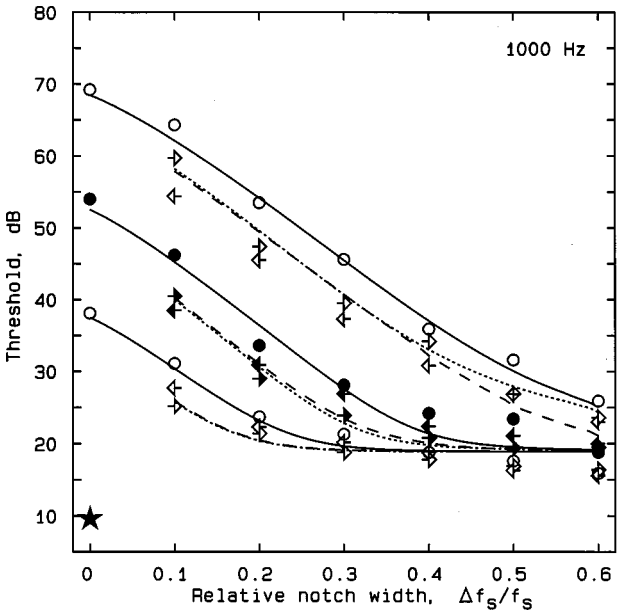


FIG. 3. As in Fig. 2, but for a signal frequency of 1000 Hz.

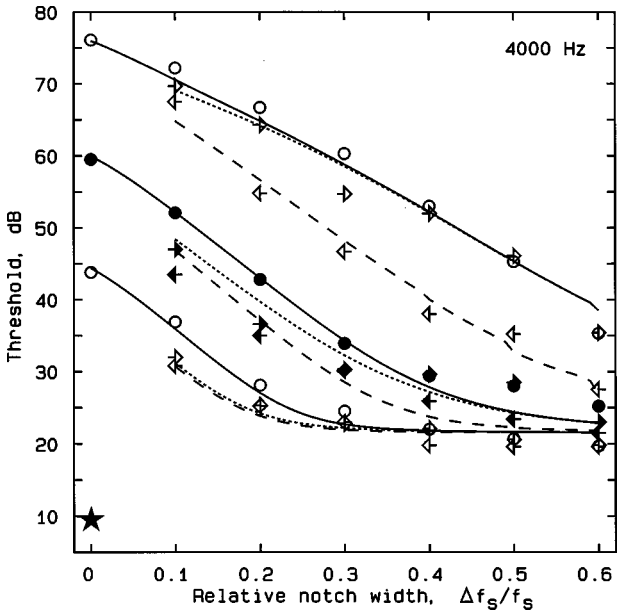


FIG. 5. As in Fig. 2, but for a signal frequency of 4000 Hz.

left-pointing arrows tend to fall below the right-pointing arrows, especially for the highest noise level (open symbols at top). This indicates that the auditory filter becomes increasingly asymmetric, with a shallower lower skirt, as the level increases. For the symmetric-notch conditions, the decrease in threshold as $\Delta f_s/f_s$ is increased from 0 to 0.6 is about 19 dB for the lowest noise level (and may be limited by the approach to absolute threshold) and 10 dB for the highest noise level, indicating that the sharpness of the filter decreases with increasing level.

For $f_s = 500$ Hz (Fig. 2), the change in asymmetry with level is similar to that observed at 250 Hz; the left- and right-pointing arrows overlap at the lowest noise level, but the left-pointing arrows lie below the right-pointing arrows at the highest noise level. At 500 Hz, the decrease in threshold with increasing notch width is not smaller at the highest noise level. However, this may have happened because, for the lower two noise levels, thresholds for large notch widths were limited by the approach to absolute threshold.

For signal frequencies of 1000, 2000, and 4000 Hz, the general pattern of the results was similar to that at 500 Hz, except that the decrease in threshold with increasing notch width tended to be larger than at 500 Hz, indicating a greater relative sharpness of the filters. The asymmetry at the highest noise level (left-pointing arrows below right-pointing arrows) was especially marked at 2000 and 4000 Hz.

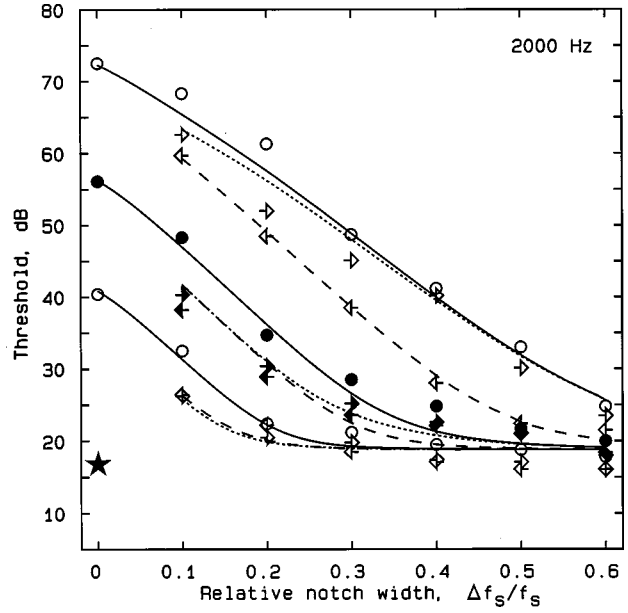


FIG. 4. As in Fig. 2, but for a signal frequency of 2000 Hz.

III. MODELING THE RESULTS

We describe here two models that can be used to analyze the data for all levels simultaneously. One model is similar to that described by Glasberg *et al.* (1999), but with some modifications. We refer to this as the masker-level-dependent (MLD) model. The other model is similar to the ''PolyFit'' method described by Rosen and co-workers (Rosen *et al.*, 1992, 1998; Baker *et al.*, 1998), but again with some modifications. We refer to this as the signal-level-dependent (SLD) model.

A. Variation of the gain of the tip filter with level

Both models are based on the assumption that the low-frequency side of the auditory filter can be represented as the sum of a tip filter and a tail filter. The gain of the tail filter is assumed to be invariant with level. The gain of the tip filter is assumed to vary with level in a similar way to the BM input–output function for a frequency close to CF. A schematic family of such input–output functions is shown in Fig. 6. These curves are based on an equation suggested by Oxenham (1995), and used subsequently by Moore *et al.* (1996) and by Oxenham and Moore (1997). The functions have shapes resembling those measured on the BM in animals (Robles *et al.*, 1986; Ruggero, 1992; Ruggero *et al.*, 1997). The functions are nearly linear for very low sound levels

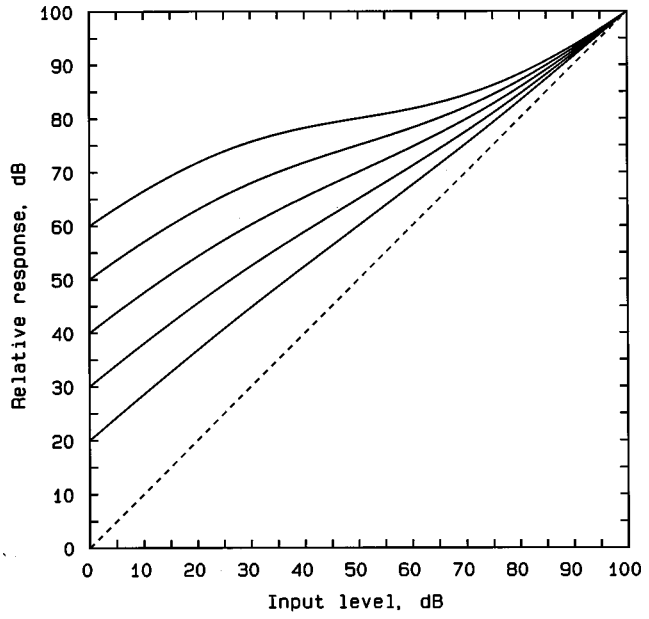


FIG. 6. Schematic input–output functions on the BM for values of the maximum gain, $G_{\max}$, ranging from 20 to 60 dB.

(i.e., the slopes are close to 1 on a dB versus dB scale), but are compressive (with slopes less than 1) for mid-range levels. The curves have been drawn so that an input level of 100 dB leads to a response of 100 dB, based on the assumption that the gain provided by the active mechanism becomes negligible at very high sound levels. Each curve is defined by the gain, G_{dB}, as a function of the input level, L, also given in dB:

$$G_{\mathrm{dB}} = 0.9L + A + B(1 - (1/(1 + \exp(-0.05(L - 50))))) - L, \tag{1}$$

where A and B are constants for a given curve. The output level is simply equal to the input level plus the gain. The constant 0.9 defines the slope of the input–output function at very low and very high levels. The constant -0.05 determines how far the central compressive region spreads out; with this value, strong compression occurs from input levels from about 20 to 80 dB. The constant 50 determines the central position of the compressive region along the x axis. The curves vary in the maximum gain, $G_{\max}$, which is applied when $L_{\mathrm{ERB}} = 0$ dB. For the uppermost curve, the output level is 60 dB for an input level of 0 dB, so $G_{\max}$ is 60 dB. For the lowest curve, $G_{\max}$ is only 20 dB. The values of A and B are chosen so that the value of G_{dB} is zero for an input level of 100 dB:

$$A = -0.0894 G_{\max} + 10.894, \tag{2}$$

$$B = 1.1789 G_{\max} - 11.789. \tag{3}$$

It is widely believed that the active mechanism provides more gain at medium and high frequencies than at low frequencies (Yates, 1990, 1995), so one aim of our modeling is to find the input–output function that gives the best fit to the data at each center frequency. Thus, the value of $G_{\max}$ is a free parameter when the model is fitted to the data.

B. Characterizing the auditory filter shape as a function of level

The upper side of the filter is modeled as a single rounded-exponential function:

$$W(g) = (1 + p_u g) \exp(-p_u g), \tag{4}$$

where g is normalized deviation from the center frequency (deviation divided by the center frequency) and p_u is a parameter determining the sharpness of the filter. Expressed in this way, the filter has unity gain at its tip. The lower side of the filter is modeled as the sum of two rounded-exponential functions, one for the tip and one for the tail:

$$W(g) = [G_{\mathrm{lin}}(1 + p_l g) \exp(-p_l g) + (1 + tg) \exp(-tg)]/(G_{\mathrm{lin}} + 1). \tag{5}$$

The term $(G_{\mathrm{lin}} + 1)$ makes the filter gain unity at its tip, which is convenient for modeling purposes, as it equates the gain at the tip for the lower and upper halves. This equation is similar to the roex(p,w,t) model proposed by Patterson *et al.* (1982), except that they used a parameter w to represent the relative gain of the tip and tail filters; the tip filter had a gain $(1-w)$ and the tail filter had a gain w. Hence, $w = 1/(G_{\mathrm{lin}} + 1)$ and $G_{\mathrm{lin}} = (1-w)/w$. The value of G_{lin} specifies the gain of the tip filter relative to that of the tail filter. Expressed in dB, it is assumed to vary with level as defined by Eq. (1).

For the MLD model, following Glasberg *et al.* (1999), the input level, L, was specified as the noise level per ERB. The ERB here is that around the frequency components of interest. For example, suppose that we wish to determine the response of an auditory filter centered at 2000 Hz to the noise components centered around 1000 Hz. The input level per ERB is obtained by summing the powers of all components in the stimulus over the appropriate frequency range around 1000 Hz (roughly from 935 to 1065 Hz; the exact range depends on the current parameter values assumed in the model, as described below). It is this input level that is used to determine the gain of the tip filter, and hence the response of the filter to the components around 1000 Hz. In practice, the level per ERB is calculated using a rounded-exponential weighting function rather than a rectangular function. The level per ERB is similar to the ''critical band level'' described by Zwicker and Fastl (1990) and used by them for the calculation of excitation patterns, except that they used a rectangular weighting function.

For the SLD model, following Rosen and co-workers (Rosen *et al.*, 1992, 1998; Baker *et al.*, 1998), the effective input level, L, was assumed to be equal to the level of the signal at threshold. Note that, in the formulation given here, this is not the same as the filter output level, since the gain of the tip filter is assumed to vary with L; the output of the tip filter is equal to $L + G_{\mathrm{dB}}$.

Several variations of the above models were tried in which the parameters p_u, p_l, and t were assumed to be either invariant with level or were polynomial functions of level. As would be expected, the data were fitted better when more free parameters were assumed (see below for details of the fitting method). In particular, the fits were somewhat bet-

ter when p_u was made a function of level. However, the changes in p_u with level varied across frequency in an inconsistent way. We found that reasonably good fits to the data could be obtained using the following assumptions: The values of p_l, p_u, and t are level invariant, and each is described by one parameter. We also assessed the effect of having a free parameter corresponding to an ''offset'' of the gain of the tip filter relative to the tail filter. For example, at very high levels, when G_{dB} becomes zero, there might still be a difference in gain between the tip filter and the tail filter. We found that the data were fitted better when such an offset was included, but that the value of the offset did not vary systematically with center frequency. Therefore, we averaged the value of the offset across frequency, and used the value obtained as a fixed offset for all center frequencies. This offset value was −8.2 dB (corresponding to a factor of 0.15 in linear power units), meaning that, at very high levels, the gain of the tip filter would be 8.2 dB less than that of the tail filter. This implies that, at high levels, the response is dominated by the tail filter.

Overall, for the model described below, the filter shape was defined by four parameters: p_l, p_u, t, and $G_{\max}$. The absolute threshold is modeled by an internal noise, and is defined by one parameter. There is also a parameter, K, representing the signal-to-masker ratio at the output of the auditory filter required to reach masked threshold; K is assumed to be level independent. This gives a total of six parameters to be varied during the fitting procedure described below.

C. Using the MLD model to derive auditory filter shapes from notched-noise data

The general procedure for deriving auditory filter shapes using the MLD model is as follows:

(1) The effective spectrum reaching the cochlea is calculated for each notch width, taking into account the response of the headphone at the eardrum and the transfer through the middle ear. The latter is assumed to have the form specified in Moore *et al.* (1997). For our UEN, the effective spectrum reaching the cochlea has a constant level per ERB within the passbands of the noise. Thus, for the MLD model, this first stage of the procedure can be eliminated and the nominal level per ERB can be used directly as input to the model.
(2) Reasonable starting values are chosen for the six parameters.
(3) For each notch width, excitation patterns are calculated for the noise alone and the noise plus the signal at a specified level. The excitation pattern is defined as the output of each auditory filter as a function of filter center frequency (Moore and Glasberg, 1983; Glasberg and Moore, 1990). The excitation patterns include the effect of the assumed internal noise related to the absolute threshold.
(4) The maximum difference in excitation level between the two excitation patterns (noise alone and noise plus signal), ΔL, is calculated. The signal level is adjusted until the value of ΔL corresponds to a fixed criterion value, C. This gives a predicted signal threshold for that notch width. The value of C corresponds to the ''efficiency parameter'' K used in earlier analyses of notched-noise data (Patterson and Moore, 1986), except that K is expressed as the signal-to-noise ratio at the output of the auditory filter required to achieve threshold, while C represents the change in excitation level produced by adding the signal to the noise.
(5) This is repeated for each notch width. The value of C is adjusted so that the mean of the predicted thresholds equals the mean of the obtained thresholds. This requires an iterative procedure, i.e., step (4) has to be repeated for each value of C, until the criterion in step (5) is satisfied.
(6) The sum of the squared deviations between the obtained and predicted thresholds (in dB) is calculated.
(7) The resulting sum of squares is entered into a multidimensional minimization routine to find the values of the six parameters that give the best fit to the data, i.e., that minimize the sum of squares.
(8) The best-fitting parameters define the filter shape for any given input level.

TABLE I. The goodness of fit for each model at each frequency, specified by the root-mean-square (rms) deviation of the data from the fitted values, in dB.

Frequency (Hz)	MLD	SLD	SLD($w$lin)
250	1.66	1.15	1.36
500	1.33	1.29	1.25
1000	1.68	1.57	1.59
2000	1.94	1.43	1.43
4000	2.36	1.36	1.38

D. Using the SLD model to derive auditory filter shapes from notched-noise data

To apply the SLD model, we used a modification of the ''PolyFit'' procedure described by Rosen and co-workers (Rosen *et al.*, 1992, 1998; Baker *et al.*, 1998). The reader is referred to their papers for details. In their procedure, the parameter w was a polynomial function of level. In our modification, the gain of the tip filter relative to the tail filter was assumed to have the form specified by Eq. (1) (but with the fixed offset of −8.2 dB). Otherwise, the procedure was exactly as described by Rosen *et al.* (1998).

E. Results of applying the models

The goodness of fit for each model at each frequency is specified in Table I by the root-mean-square (rms) deviation of the data from the fitted values. Overall, the data were fitted markedly better by the SLD model than by the MLD model. This is consistent with the finding of Rosen and coworkers (Rosen *et al.*, 1992, 1998; Baker *et al.*, 1998) that models in which the filter shape depends on the signal level fit the data better than models in which the filter shape depends on the masker level. For comparison, the table also shows the goodness of fit obtained with the signal-dependent PolyFit model when w was assumed to be a linear function of level, defined by two free parameters. As for the SLD model, the parameters p_l, t, and p_u were assumed to be level

TABLE II. Parameters defining the shapes of the filters derived using the SLD model for each center frequency. Also shown is the absolute threshold, Abs (dB) estimated from the model, and the value of the parameter K (dB), which represents the signal-to-masker ratio required at the output of the auditory filter to reach masked threshold.

Center frequency (Hz)	p_l	t	$G_{\max}$ (dB)	p_u	Abs	K (dB)
250	18.5	6.5	22.9	12.9	16.0	1.9
500	29.2	12.1	24.5	16.5	22.0	−2.3
1000	29.5	8.6	47.5	22.1	18.9	−3.2
2000	39.2	13.8	35.2	28.3	18.8	−0.8
4000	34.0	12.0	34.6	26.3	21.4	3.4

independent. The value of w (in dB) was not allowed to become positive, a constraint also applied in the analyses of Rosen and co-workers (a positive value of w would mean that the power at the output of the tip filter was subtracted from the power at the output of the tail filter, which is clearly unrealistic). This version of the PolyFit model, called the SLD($w$lin) model, therefore has seven free parameters, whereas the SLD model has six (excluding the fixed offset of −8.2 dB). The two models fit the data almost equally well, except at 250 Hz where the SLD model fits better, in spite of the fact that the SLD($w$lin) model has one more free parameter.

The parameters defining the derived filters obtained with the SLD model are given in Table II. It can be seen that the values of p_l, p_u, and $G_{\max}$ all increase with increasing center frequency from 250 to 1000 Hz, but do not show a consistent trend with further increases in center frequency. This is consistent with both physiological data (Yates, 1990, 1995; Neely, 1993; Rhode and Cooper, 1996) and psychoacoustic data (Hicks and Bacon, 1999; Moore *et al.*, 1999; Plack and Oxenham, 2000) suggesting that the active mechanism in the cochlea plays a stronger role at medium and high frequencies than at low frequencies, although physiological measurements of BM motion (Cooper and Rhode, 1995; Ruggero *et al.*, 1997) exist mainly for low and high CFs, and little is known about basilar-membrane responses to mid-range frequencies.

Table II also shows the absolute thresholds estimated by the SLD model. For the signal frequencies of 250 and 500 Hz, these thresholds were within 0.7 dB of the measured absolute thresholds (after allowing for the transmission of the signal through the headphone and the outer and middle ear). However, for the higher signal frequencies, the absolute thresholds estimated by the SLD model were 2–11 dB higher than the measured absolute thresholds. This discrepancy appears to have two sources. First, for the three higher signal frequencies, the predicted thresholds reached an asymptote slightly above the obtained thresholds. This may indicate that the parameters used to model the filter shape were not sufficient to account for the ''flattening'' of the thresholds at large notch widths, and the absolute threshold parameter was used to produce this flattening. As a result, thresholds for large notch widths were predicted to be the same for the symmetric and asymmetric notch conditions, whereas in fact they differed slightly. Second, the masked thresholds for the higher center frequencies flattened off at values slightly above the measured absolute thresholds. This phenomenon has been observed previously (Glasberg *et al.*, 1984), and it may reflect masking by distortion products produced by the upper band of noise (Glasberg and Moore, 1994; Moore *et al.*, 1995). The first source of discrepancy may be related to the second; if the thresholds flatten off at large notch widths because of distortion products, then the flattening should not be regarded as arising from the filter shape *per se*.

When the gain of the tip filter is large relative to the gain of the tail filter (i.e., when G_{lin} is large), the ERBs of the filters are approximately equal to $(2/p_l + 2/p_u) f_s$ (Patterson and Nimmo-Smith, 1980). ERBs calculated in this way are 67, 94, 157, 242, and 548 Hz, for $f_s = 250$, 500, 1000, 2000, and 4000 Hz, respectively. These values are somewhat greater than the values of 52, 79, 132, 240, and 456 Hz, respectively, proposed by Glasberg and Moore (1990) for an input level of 51 dB per ERB. The discrepancy may reflect individual differences in the frequency selectivity of the subjects, or, more likely, it may reflect differences in the model fitted to the data. Glasberg and Moore assumed that the filter shape depended on the masker level per ERB, and fitted most of the data with the roex(p,r) filter shape proposed by Patterson *et al.* (1982), which models each side of the auditory filter as a single rounded exponential with a hard dynamic range limitation (i.e., the filter becomes completely flat at large deviations from the center frequency). For the SLD model, the filter was assumed to have a sloping tail on the low-frequency side and no shallow portion on the high-frequency side. The use of the roex(p,r) model tends to lead to larger p values (and hence smaller ERBs), especially for the high-frequency side of the auditory filter (determined by parameter p_u).

Thresholds predicted by the SLD model are shown as smooth curves in Figs. 1–5. The derivation of these predictions requires an iterative procedure, as the filter shape is assumed to depend on the signal level, and the signal level at threshold cannot be predicted directly from the noise level and notch width. At each stage of the iteration, the predicted thresholds from the previous stage are used to define the parameters of the filter to generate a new set of predictions. This is repeated until the predicted thresholds stabilize. As a result, the predicted thresholds shown in Figs. 1–5 are not identical to the fitted values used to derive the goodness-of-fit measures in Table I. The fitted values are predicted from the obtained thresholds.

In general, there were few systematic deviations from the data, with the following exceptions: (1) For $f_s = 250$ Hz, the obtained thresholds for the noise levels of 50 and 65 dB/ERB tended to be slightly higher than predicted for medium notch widths (see Fig. 1). (2) For $f_s = 500$, 1000, 2000, and 4000 Hz, the predicted thresholds reached an asymptote slightly above the obtained thresholds. (3) For $f_s = 2000$ and 4000 Hz, the obtained thresholds were slightly higher than predicted at the highest noise level for symmetric notches of width 0.1 to 0.3 (see Figs. 4 and 5).

We have assumed, as did Rosen and co-workers, that the filter shape, and in particular the gain of the tip filter relative to the tail filter, is related to the signal level at threshold, L_s. As noted earlier, in the formulation given here, the signal

level is *not* the same as the filter output level, since the gain of the tip filter is assumed to vary with L_s; the output of the tip filter is equal to $L_s + G_{\mathrm{dB}}$. This means that the filter shape is not a simple function of either the input level or the output level. However, the filter shape can be thought of as being determined by the *normalized* output level, i.e., the output level when the filter is scaled to have 0-dB gain at its center frequency.

To plot representative filter shapes, we have chosen to calculate normalized output levels and gains for several fixed input levels. The steps in doing this are as follows. Assume that we wish to calculate the filter response for a sinusoidal input with frequency f_x and level L_x. (1) We pick a starting (arbitrary) value for G_{dB}. (2) The normalized filter output for this value of G_{dB} is calculated using Eq. (5) [for values of f_x below the center frequency; otherwise Eq. (4) is used, and the value of G_{dB} is irrelevant]. This gives an initial normalized output level L_o. (3) The value of L_o is used to calculate a new value of G_{dB} [using Eq. (1), but taking into account the fixed offset of -8.2 dB]. (4) Steps 2 and 3 are repeated until the values of G_{dB} and L_o become stable. The normalized gain of the filter, assuming a gain of 0 dB at the tip, is calculated as $L_o - L_x$. The ''true'' relative gain, i.e., the gain that we would expect to see on the BM, is calculated as $L_o - L_x + 10\log_{10}(G_{\mathrm{lin}} + 1)$. To generate complete filter shapes, this procedure is repeated for many values of f_x.

The filter shapes calculated in this way for $f_s = 250$ Hz are shown in Fig. 7. The input level was varied from 30 to 80 dB in 10-dB steps. The upper panel shows normalized gains while the lower panel shows ''true'' gains. In the lower panel, the gain at the tip increases progressively with decreasing input level, while the gain on the low-frequency tail is almost invariant with level, as observed for BM responses.

The filter shapes for the center frequencies of 1000 and 4000 Hz are shown in Figs. 8 and 9. In each figure, the frequency range extends from $0.4f_s$ to $1.6f_s$. Generally, the changes in shape with level are similar to those in Fig. 7, but the filters tend to become sharper with increasing center frequency, and the changes in peak gain with level are greater than at 250 Hz. For the center frequency of 1000 Hz, the normalized gains on the low-frequency side of the filter change little for input levels from 30 to 70 dB, and then change abruptly when the level is increased to 80 dB. However, the ''true'' gains in the lower panel show more progressive changes with level.

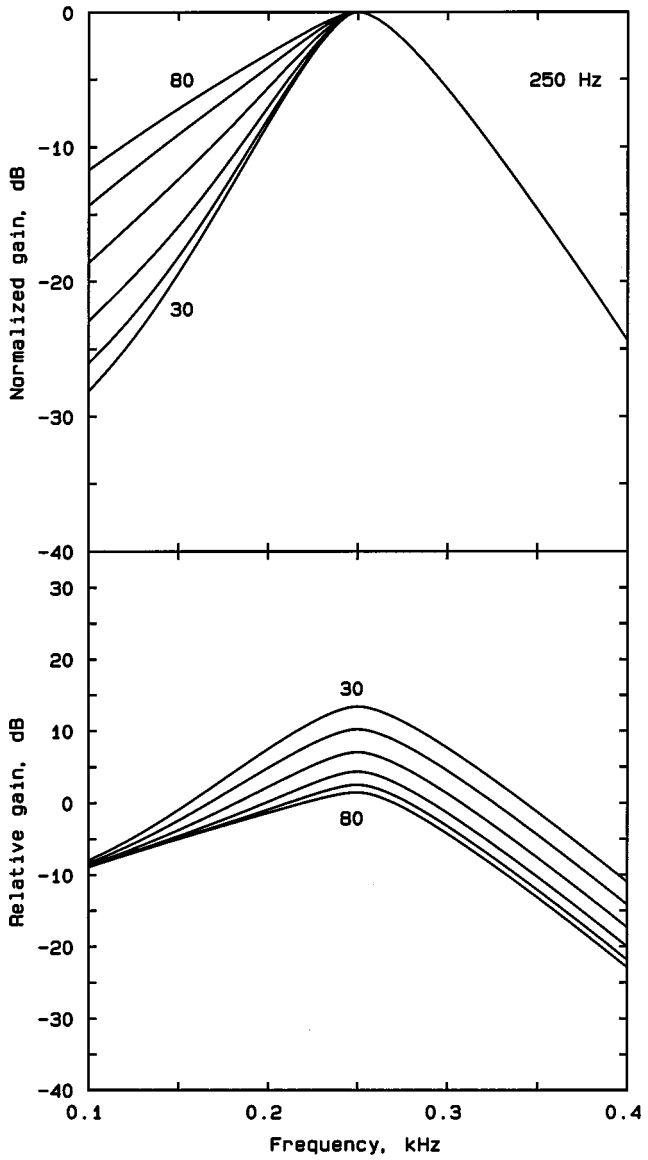


FIG. 7. Auditory filter shapes derived from the data for $f_s = 250$ Hz. Filter shapes are plotted for sinusoidal inputs with levels ranging from 30 to 80 dB in 10-dB steps. Filter responses are plotted assuming a gain of unity (0 dB) at the tips (upper panel) or in terms of filter gain (lower panel).

IV. DISCUSSION

Baker *et al.* (1998) have also presented data from notched-noise masking experiments using a wide range of center frequencies and levels; their data were analyzed using the ''PolyFit'' model, which also assumes a tip filter and a tail filter. Like us, they found a clear trend for the gain of the tip filter relative to the tail filter to increase with increasing center frequency. We have reanalyzed their data using the six-parameter SLD model and compared the results to those obtained using the seven-parameter PolyFit model described earlier [SLD(*w*lin)], in which *w* was assumed to be a linear function of level. For the SLD model, the ''offset'' in gain between the tip and tail filters was set to the same value as determined for our data (-8.2 dB). The goodness of the fits obtained with the two models is summarized in Table III. Generally, the SLD model fitted the data as well as or better than the SLD(*w*lin) model, in spite of having one less free parameter.

The values of the parameters of the filter shapes derived using the SLD model are shown in Table IV. The general pattern of the parameters is very similar to that obtained for the analysis of our own data. Specifically, the values of p_l, p_u, and G_{max} all increase with increasing center frequency from 250 to 1000 Hz, but do not show a consistent trend with further increases in center frequency. It is noteworthy that the parameter t is almost invariant with center frequency, ranging from 7.1 to 9.6. Our own data show somewhat more variability in the value of t, but with no consistent trend across frequency. Overall, the results indicate that the parameters of the tail filter do not vary markedly with frequency.

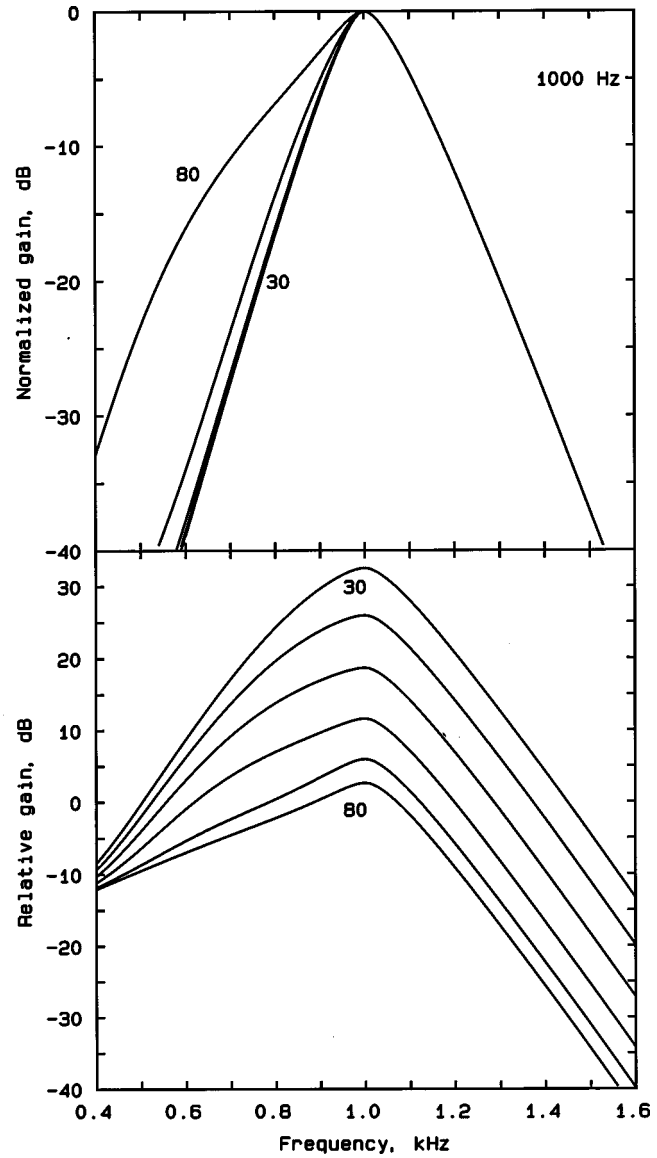


FIG. 8. As in Fig. 7, but for $f_S = 1000$ Hz. The filter shapes for input levels of 30 and 80 dB are based on extrapolation.

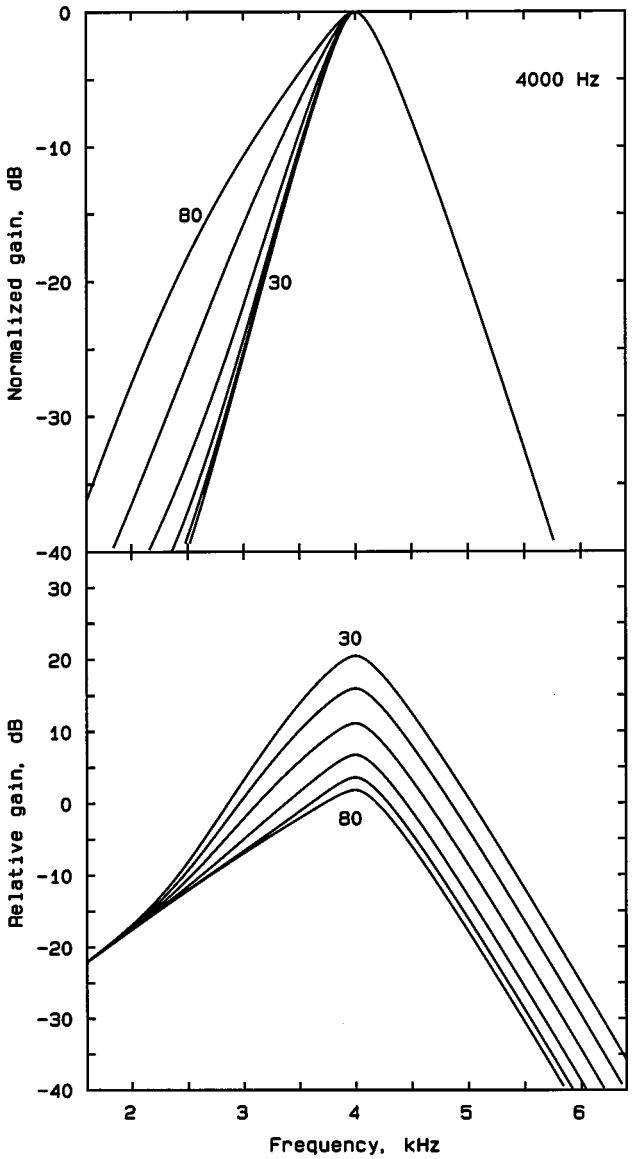


FIG. 9. As in Fig. 8, but for $f_s = 4000$ Hz.

For mid-range frequencies, Baker *et al.* found that the gain of the tip filter changed by about 20 dB when the signal level at threshold changed by 40 dB, implying a compression in the input–output function of about 2:1. They stated that ''this is similar to the degree of CF compression observed in BM motion.'' Our estimates of compression are similar. For example, if $G_{\max}$ is 40 dB (a ''representative'' value for frequencies of 1000 Hz and above), then the gain changes by 20 dB when the signal level changes from 30 to 70 dB, implying a compression of about 2:1. However, BM input–output functions for high CFs can show much more compression than this (Russell and Murugasu, 1997; Ruggero *et al.*, 1997). Also, psychoacoustical estimates of basilar-membrane compression, based on growth-of-masking functions in forward masking (Oxenham and Plack, 1997a; Moore *et al.*, 1999), suggest compression between 3:1 and 5:1 for mid-range levels.

The discrepancy between the amount of compression estimated from the notched-noise data and the more direct estimates of basilar-membrane compression may occur because simultaneous masking is partly mediated by suppression of the signal by the masker. This is especially the case when the masker is relatively intense and when masking is determined mainly by the lower frequencies in the masker (Delgutte, 1988, 1990; Moore and Vickers, 1997; Oxenham and Plack, 1997b). When the six-parameter model is fitted to the data, the characteristics of the tail filter, and the gain of the tip

TABLE III. The goodness of fit for the SLD and SLD(*w*lin) models to the data of Baker *et al.* (1998), specified by the root-mean-square (rms) deviation of the data from the fitted values, in dB.

Frequency (Hz)	SLD	SLD(*w*lin)
250	1.16	1.60
500	1.24	1.22
1000	1.45	1.49
2000	0.80	0.85
3000	1.06	1.07
4000	1.56	1.54
6000	1.62	1.66

TABLE IV. Parameters defining the shapes of the filters derived using the SLD model applied to the data of Baker *et al.* (1998) for each center frequency. Also shown is the absolute threshold, Abs (dB) estimated from the model, and the value of the parameter K (dB), which represents the signal-to-masker ratio required at the output of the auditory filter to reach masked threshold.

Center frequency (Hz)	p_l	t	G_{max} (dB)	p_u	Abs	K (dB)
250	17.2	7.1	39.2	19.5	20.8	1.2
500	22.9	7.7	39.7	24.8	17.9	−1.6
1000	36.3	8.9	43.4	31.6	11.1	−1.4
2000	33.7	8.7	44.1	27.6	15.3	−2.3
3000	31.7	7.5	49.0	26.8	19.4	−1.6
4000	45.5	9.6	47.7	28.9	21.2	0.3
6000	45.4	9.4	50.2	26.5	18.8	0.6

filter relative to the tail filter, are mainly determined by the data for large notch widths and high masker levels. These are precisely the conditions where suppression probably plays a role. The thresholds in these conditions probably do not depend solely on the spread of masker excitation to the signal frequency; they also depend on suppression of the signal by the masker. Thus, the derived filters represent the combined effects of spread of excitation and suppression. If the effects of spread of excitation could be measured alone (for example, by using forward masking), then the derived filters would probably be sharper and the gain of the tip filter would change more with level.

Like Rosen and his co-workers, we found that the data were fitted better by a model in which the filter parameters were assumed to depend on the signal level (SLD model) than by a model in which the parameters depended on the masker level (MLD model). Note that, for the MLD model, Rosen and co-workers assumed that the masker spectrum level was the appropriate variable, whereas we assumed that the masker level per ERB was the appropriate variable. However, the outcome was similar for these two sets of assumptions.

In their analyses, Rosen and co-workers assumed that the signal level at masked threshold was closely related to the magnitude of the output of the auditory filter; the output was assumed to be equal to the signal level at threshold plus a constant (in dB). Expressed in linear power, the output of the filter was assumed to be directly proportional to the signal power. Thus, they characterized their results as indicating that the filter shape depends on its output level. In our analyses, we have assumed that the gain of the tip filter depends on the signal level. In this case, the output power of the filter is not directly proportional to the signal level. Thus the results should not be taken as indicating that the filter shape depends simply on the power at its output. Rather, the filter shape depends on the power at the output divided by the gain (or, in dB terms, the level at the output minus the gain). This dependence may arise because filtering on the BM involves a feedback mechanism, which means that it is unrealistic to characterize the shape as dependent only on the input level or only on the output level. Another possibility is that the tail and tip filters effectively act in series, with the output of the tail filter providing the input to the tip filter. The gain of the tip filter might depend on the level at the output of the tail filter. We have yet to explore this possibility in detail.

In summary, the present approach to the derivation of auditory filter shapes from notched-noise data is more closely related than previous approaches to the physiological properties of the filtering measured on the BM. The approach is clearly over-simplified, and needs further refinement. Nevertheless, auditory filter shapes derived using the method do vary with level in a way that closely resembles the level variation of tuning on the BM. This encourages us to believe that the basic properties of the auditory filters are determined in the cochlea. It should be remembered, however, that measures of frequency selectivity derived from simultaneous-masking data may be influenced by suppression.

ACKNOWLEDGMENTS

This work was supported by the Medical Research Council (UK). We thank Stuart Rosen and Richard Baker for supplying us with the PolyFit software and with their raw data; Stuart Rosen, Michael Stone, and Roy Patterson for helpful discussions; and Marjorie Leek and Stuart Rosen for very helpful comments on an earlier version of this paper.

[1]A DOS program for converting between frequency, ERB width and ERB number is available at http://hearing.psychol.cam.ac.uk/.

Baker, R. J., Rosen, S., and Darling, A. M. (**1998**). ''An efficient characterisation of human auditory filtering across level and frequency that is also physiologically reasonable,'' in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London).

Burkhard, M. D., and Sachs, R. M. (**1975**). ''Anthropometric manikin for acoustic research,'' J. Acoust. Soc. Am. **58**, 214–222.

Cooper, N. P., and Rhode, W. S. (**1995**). ''Nonlinear mechanics at the apex of the guinea-pig cochlea,'' Hear. Res. **82**, 225–243.

Delgutte, B. (**1988**). ''Physiological mechanisms of masking,'' in *Basic Issues in Hearing*, edited by H. Duifhuis, J. W. Horst, and H. P. Wit (Academic, London).

Delgutte, B. (**1990**). ''Physiological mechanisms of psychophysical masking: Observations from auditory-nerve fibers,'' J. Acoust. Soc. Am. **87**, 791–809.

Evans, E. F., Pratt, S. R., and Cooper, N. P. (**1989**). ''Correspondence between behavioural and physiological frequency selectivity in the guinea pig,'' Br. J. Audiol. **23**, 151–152.

Fletcher, H. (**1940**). ''Auditory patterns,'' Rev. Mod. Phys. **12**, 47–65.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hear. Res. **47**, 103–138.

Glasberg, B. R., and Moore, B. C. J. (**1994**). ''Growth-of-masking functions for several types of maskers,'' J. Acoust. Soc. Am. **96**, 134–144.

Glasberg, B. R., Moore, B. C. J., and Stone, M. A. (**1999**). ''Modelling changes in frequency selectivity with level,'' in *Psychophysics, Physiology and Models of Hearing*, edited by T. Dau, V. Hohmann, and B. Kollmeier (World Scientific, Singapore).

Glasberg, B. R., Moore, B. C. J., Patterson, R. D., and Nimmo-Smith, I. (**1984**). ''Dynamic range and asymmetry of the auditory filter,'' J. Acoust. Soc. Am. **76**, 419–427.

Hicks, M. L., and Bacon, S. P. (**1999**). ''Psychophysical measures of auditory nonlinearities as a function of frequency in individuals with normal hearing,'' J. Acoust. Soc. Am. **105**, 326–338.

Lutfi, R. A., and Patterson, R. D. (**1984**). ''On the growth of masking asymmetry with stimulus intensity,'' J. Acoust. Soc. Am. **76**, 739–745.

Moore, B. C. J., (**1986**). ''Parallels between frequency selectivity measured psychophysically and in cochlear mechanics,'' Scand. Audiol. Suppl. **25**, 139–152.

Moore, B. C.. J., and Glasberg, B. R. (**1983**). ''Suggested formulae for calculating auditory-filter bandwidths and excitation patterns,'' J. Acoust. Soc. Am. **74**, 750–753.

Moore, B. C. J., and Glasberg, B. R. (**1987**). ''Formulae describing frequency selectivity as a function of frequency and level and their use in calculating excitation patterns,'' Hear. Res. **28**, 209–225.
Moore, B. C. J., and Vickers, D. A. (**1997**). ''The role of spread of excitation and suppression in simultaneous masking,'' J. Acoust. Soc. Am. **102**, 2284–2290.
Moore, B. C. J., Glasberg, B. R., and Baer, T. (**1997**). ''A model for the prediction of thresholds, loudness and partial loudness,'' J. Audio Eng. Soc. **45**, 224–240.
Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1996**). ''Detection of decrements and increments in sinusoids at high overall levels,'' J. Acoust. Soc. Am. **99**, 3669–3677.
Moore, B. C. J., Vickers, D. A., Plack, C. J., and Oxenham, A. J. (**1999**). ''Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism,'' J. Acoust. Soc. Am. **106**, 2761–2778.
Moore, B. C. J., Glasberg, B. R., van der Heidjen, M., Houtsma, A. J. M., and Kohlrausch, A. (**1995**). ''Comparison of auditory filter shapes obtained with notched-noise and noise-tone maskers,'' J. Acoust. Soc. Am. **97**, 1175–1182.
Neely, S. T. (**1993**). ''A model of cochlear mechanics with outer hair cell motility,'' J. Acoust. Soc. Am. **94**, 137–146.
Oxenham, A. J. (**1995**). ''Psychophysical consequences of peripheral auditory nonlinearity,'' Ph.D. Thesis, Cambridge.
Oxenham, A. J., and Moore, B. C. J. (**1997**). ''Modeling the effects of peripheral nonlinearity in listeners with normal and impaired hearing,'' in *Modeling Sensorineural Hearing Loss*, edited by W. Jesteadt (Erlbaum, Hillsdale, NJ).
Oxenham, A. J., and Plack, C. J. (**1997a**). ''A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing,'' J. Acoust. Soc. Am. **101**, 3666–3675.
Oxenham, A. J., and Plack, C. J. (**1997b**). ''Suppression and the upward spread of masking,'' J. Acoust. Soc. Am. **104**, 3500–3510.
Patterson, R. D., and Moore, B. C. J. (**1986**). ''Auditory filters and excitation patterns as representations of frequency resolution,'' in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London).
Patterson, R. D., and Nimmo-Smith, I. (**1980**). ''Off-frequency listening and auditory filter asymmetry,'' J. Acoust. Soc. Am. **67**, 229–245.
Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (**1982**). ''The deterioration of hearing with age: frequency selectivity, the critical ratio, the audiogram, and speech threshold,'' J. Acoust. Soc. Am. **72**, 1788–1803.
Plack, C. J., and Oxenham, A. J. (**2000**). ''Basilar-membrane nonlinearity estimated by pulsation threshold,'' J. Acoust. Soc. Am. **107**, 501–507.
Puria, S., Rosowski, J. J., and Peake, W. T. (**1997**). ''Sound-pressure measurements in the cochlear vestibule of human-cadaver ears,'' J. Acoust. Soc. Am. **101**, 2754–2770.
Rhode, W. S., and Cooper, N. P. (**1996**). ''Nonlinear mechanics in the apical turn of the chinchilla cochlea *in vivo*,'' Aud. Neurosci. **3**, 101–121.
Rhode, W. S., and Robles, L. (**1974**). ''Evidence from Mössbauer experiments for non-linear vibration in the cochlea,'' J. Acoust. Soc. Am. **55**, 588–596.
Robles, L., Ruggero, M. A., and Rich, N. C. (**1986**). ''Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input-output functions, tuning curves, and response phases,'' J. Acoust. Soc. Am. **80**, 1364–1374.
Rosen, S., Baker, R. J., and Darling, A. (**1998**). ''Auditory filter nonlinearity at 2 kHz in normal hearing listeners,'' J. Acoust. Soc. Am. **103**, 2539–2550.
Rosen, S., Baker, R. J., and Kramer, S. (**1992**). ''Characterizing changes in auditory filter bandwidth as a function of level,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, K. Horner, and L. Demany (Pergamon, Oxford).
Ruggero, M. A. (**1992**). ''Responses to sound of the basilar membrane of the mammalian cochlea,'' Curr. Opin. Neurobiol. **2**, 449–456.
Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (**1997**). ''Basilar-membrane responses to tones at the base of the chinchilla cochlea,'' J. Acoust. Soc. Am. **101**, 2151–2163.
Russell, I. J., and Murugasu, E. (**1997**). ''Medial efferent inhibition suppresses basilar membrane responses to near characteristic frequency tones of moderate to high intensities,'' J. Acoust. Soc. Am. **102**, 1734–1738.
Sellick, P. M., Patuzzi, R., and Johnstone, B. M. (**1982**). ''Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique,'' J. Acoust. Soc. Am. **72**, 131–141.
Yates, G. K. (**1990**). ''Basilar membrane nonlinearity and its influence on auditory nerve rate-intensity functions,'' Hear. Res. **50**, 145–162.
Yates, G. K. (**1995**). ''Cochlear structure and function,'' in *Hearing*, edited by B. C. J. Moore (Academic, San Diego).
Zwicker, E. (**1970**). ''Masking and psychological excitation as consequences of the ear's frequency analysis,'' in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Leiden).
Zwicker, E., and Fastl, H. (**1990**). *Psychoacoustics—Facts and Models* (Springer-Verlag, Berlin).

Frequency modulation detection interference produced by asynchronous and nonsimultaneous interferers

Hedwig Gockel[a] and Robert P. Carlyon
MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF, United Kingdom

(Received 23 February 2000; revised 5 June 2000; accepted 16 August 2000)

The effect of asynchronous and nonsimultaneous interferers on detection of sinusoidal frequency modulation (FM) was compared with the effect of a synchronous interferer. In a two-interval, two-alternative forced-choice (2I-2AFC) adaptive procedure, listeners had to detect FM with a modulation frequency of 15 Hz, imposed on a 1-kHz sinusoidal carrier (the target). The 200-ms target was presented either alone (baseline condition), or with an interferer whose timing relative to the target was varied. The interferer was a 2.3-kHz sinusoidal carrier which was also frequency modulated at a rate of 15 Hz. Experiment one showed that thresholds for detection of FM increased significantly, both with a synchronous FM interferer, and also with asynchronous interferers (starting 200 ms before and stopping 200 ms after the target). Moreover, ''gapped'' interferers that were turned off during presentation of the target (presented for 200 ms before and for 200 ms after the target but not simultaneously) produced the same significant increase in thresholds as an asynchronous interferer that was not interrupted. In contrast, thresholds were not affected by the presence of a gapped unmodulated sinusoidal interferer. Experiment two showed that increasing the duration of the silent gap (centered on presentation of the target) between FM interferers from 200 to 600 ms did not abolish the interference. Thus nonsimultaneous FM interferers produced frequency modulation detection interference (FMDI) even when the silent gap between the interferers and target clearly led to the interferers and target being perceived as separate auditory objects. A possible explanation for the findings is the existence of an asymmetry in perception of steady and modulated sounds, as recently proposed by Cusack and Carlyon [Br. J. Audiol. **34.2**, 112 (2000)]. Alternative explanations in terms of ringing in a hypothetical modulation filter bank and adaptation seem unlikely. [S0001-4966(00)03811-X]

PACS numbers: 43.66.Mk, 43.66.Lj, 43.66.Ba [DWG]

I. INTRODUCTION

The detection of amplitude or frequency modulation of a target tone can be substantially impaired by the presence of modulation imposed on a carrier frequency remote from that of the target tone. This phenomenon, called modulation detection interference (MDI), has been the subject of many investigations. Most of this research has concentrated on the effects of amplitude modulated tones on the detection of either the presence of amplitude modulation (AM) or a change in the depth of amplitude modulation of a target tone (Yost and Sheft, 1989; Yost *et al.*, 1989; Yost and Sheft, 1990; Hall and Grose, 1991; Moore and Jorasz, 1992; Moore and Shailer, 1992; Bacon and Konrad, 1993; Bacon and Moore, 1993; Bacon and Opie, 1994; Bacon *et al.*, 1995; Mendoza *et al.*, 1995). Both of these will be abbreviated as AMDI in the present paper. In contrast, only a few studies have investigated the effect of frequency modulated tones on the detection of either the presence of frequency modulation (FM) or a change in the depth of frequency modulation of a target (Wilson *et al.*, 1990; Moore *et al.*, 1991; Carlyon, 1994; Lyzenga and Carlyon, 1999), both of which will be abbreviated as FMDI.

MDI decreases with increasing separation of the carrier frequencies (f_c) of the interfering modulated tones (interferers) and the target tone, but substantial MDI is still observed for interferers separated from the target by 2 to 3 octaves (Yost *et al.*, 1989; Wilson *et al.*, 1990; Moore and Jorasz, 1992; Bacon and Konrad, 1993; Bacon and Moore, 1993; Mendoza *et al.*, 1995). Thus across-channel processes have been assumed to play a major role in MDI. The observation that interferers presented contralaterally to the target also produce AMDI (Yost and Sheft, 1990; Bacon and Opie, 1994; Mendoza *et al.*, 1995), supports an interpretation in terms of central auditory processes. One specific explanation of AMDI based on across-channel processes, was put forward by Yost and Sheft (1989). They suggested that target and interferer (which were presented simultaneously) are perceptually grouped together due to the common modulation. This perceptual fusion would make it difficult to hear out the modulation of the target sound. However, MDI is tuned only very coarsely to rate of modulation (Yost *et al.*, 1989; Wilson *et al.*, 1990; Moore *et al.*, 1991; Bacon *et al.*, 1995). As has been argued by Moore *et al.* (1991), and Moore and Shailer (1992), a much sharper tuning would be expected from an auditory grouping perspective.

A considerable amount of effort has been put into clarifying the role of auditory grouping in MDI. The general approach has been to investigate the influence on MDI of parameters which influence perceptual grouping of simultaneous components, i.e., concurrent sound segregation.

[a]Present address: CNBH, Department of Physiology, University of Cambridge, Downing Street, Cambridge CB2 3EG, U.K. Electronic mail: hedwig.gockel@mrc-cbu.cam.ac.uk

Among the parameters which have been investigated are onset and offset asynchrony, duration, relative ear of presentation, harmonicity, and spectral proximity. The results show: (a) With increasing onset and/or offset asynchrony of the target and modulated interferer, the amount of AMDI decreases (Hall and Grose, 1991; Moore and Jorasz, 1992; Moore and Shailer, 1992), but some AMDI is still present even for continuously presented interferers (Hall and Grose, 1991; Mendoza *et al.*, 1995). (b) AMDI decreases with increasing duration of the (synchronously gated) target and interferer (Moore and Shailer, 1992). (c) As mentioned above, dichotic presentation of the target and interferer did produce AMDI, although the effect was smaller for contralateral than for ipsilateral presentation of the target and interferer. (d) The influence of harmonicity was somewhat mixed. On one hand, Bacon and Moore (1993) found no more AMDI for interferers which were harmonically related to the target than for nonharmonic interferers equally spaced on an ERB scale (gated synchronously in both cases). On the other hand, Hall and Grose (1991) found a decrease in AMDI when the interferer was accompanied by several harmonics of a complex tone which were harmonically related and spectrally proximal to the interferer. However, this ''capturing'' of the interferer into a complex tone depended on the extra harmonics and the interferer being gated on and off simultaneously. (e) Mendoza *et al.* (1995) found an increase in AMDI with increase in frequency proximity of target and interferer even when the interferer was presented contralaterally to the target. Thus even in the absence of within-channel effects spectral proximity plays a role.

Overall, these findings show that changing the stimuli in a way which enhances the segregation of concurrent sounds also tends to reduce the amount of MDI. This strongly suggests that perceptual grouping is important for MDI. However, introducing onset and offset asynchrony, one of the most effective factors in concurrent sound segregation (Gockel, 2000), does not completely eliminate the interference. In contrast, an onset asynchrony of 320 ms is sufficient to completely eliminate the influence of a slightly mistuned harmonic on the overall pitch of a complex tone (Darwin and Ciocca, 1992). This raises the question of whether those parts of the interferer which are nonoverlapping in time with the target might themselves impair performance in detection of modulation of the target tone. To our knowledge, in all published research on the topic of MDI, the targets have always been accompanied by interferers over their whole time course of presentation, thus treating MDI as a problem of concurrent sound segregation. In the present experiments, we investigated the effect of nonsimultaneous nonoverlapping interferers on frequency modulation detection, thus investigating MDI in the context of sequential sounds.

In our first experiment, we show that the interference was as large for interferers which did not overlap in time with the target as for asynchronous modulated interferers which did overlap in time. Thus that part of the interferer which was presented simultaneously with the target did not increase FMDI once the onset and offset asynchrony was introduced. In our second experiment, the duration of the silent gap between asynchronous nonoverlapping interferers and the target was increased from 0 up to 200 ms, to enhance the perception of distinct auditory events following each other in time. The results showed that the effect of the asynchronous nonoverlapping interferers was almost unaffected by the duration of the silent gap between interferers and target.

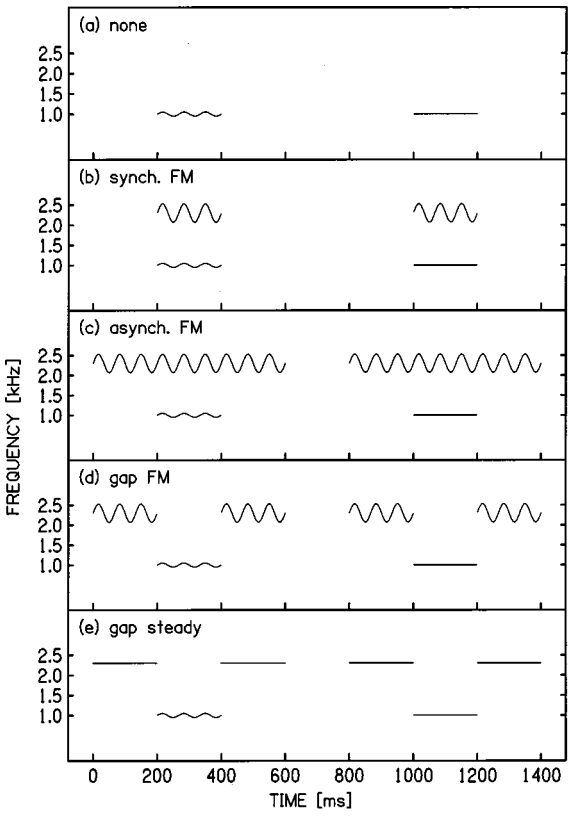


FIG. 1. Schematic spectrograms of stimuli presented over the course of one 2AFC trial in the different conditions: (a) baseline, (b) synchronous FM, (c) asynchronous FM, (d) gap FM, (e) gap steady. Note that for illustrative purposes only (i) an extreme FM depth has been chosen for the target, and (ii) the signal interval is always the first interval.

II. EXPERIMENT 1: ASYNCHRONOUS AND GAPPED INTERFERERS

A. Stimuli

Listeners had to detect sinusoidal FM with a modulation frequency (f_m) of 15 Hz, imposed on a 200-ms target with a carrier frequency (f_c) of 1 kHz. The target was either presented alone (condition ''none''), or was accompanied by an interferer. The carrier frequency of the interferer was 2.3 kHz. The interferer was either modulated with f_m of 15 Hz, with an FM depth of 10% of the carrier frequency (resulting in a peak deviation from f_c of 230 Hz) or unmodulated. Schematic spectrograms of the stimuli are shown in Fig. 1. Condition ''none'' [Fig. 1(a)] provided the baseline against which thresholds in the other conditions were compared. In condition ''synchronous FM'' [Fig. 1(b)], the interferer was presented synchronously with the target and was modulated in phase with the target. This condition was included to check that the present subjects would show the well-established MDI effect for simultaneous and synchronous interferers. In condition ''asynchronous FM'' [Fig. 1(c)], the

600-ms interferer was modulated, and started 200 ms before and stopped 200 ms after the target. This condition was run to see if the reduction of AMDI found in the past with asynchronous gating would also occur for FMDI. In condition ''gap FM'' [Fig. 1(d)], the interferer was modulated and presented for 200 ms immediately before and after the target (stopping when the target started and starting when the target stopped). Comparison of thresholds in this condition with those in condition asynchronous FM would indicate the contribution to FMDI of the part of the interferer presented simultaneously with the target (for asynchronous interferers). In condition ''gap steady'' [Fig. 1(e)], the interferer was unmodulated and was presented for 200 ms before and after (but not simultaneously with) the target. This condition was introduced to check on a possible disruptive effect of alternating the target with another nontarget sound. If thresholds were greater in condition ''gap FM'' than in condition ''gap steady,'' this would indicate that the FM of nonoverlapping signals *per se* contributes to FMDI. If thresholds in condition ''gap steady'' were equal to thresholds in condition ''gap FM,'' but greater than in condition ''none,'' this would indicate that alternating the target and the interferer leads to impairment, not the FM of the interferer *per se*. The starting phase of the modulator was always 0° (sine phase). Over the stimulus duration of 200 ms (including 10-ms raised-cosine onset and offset ramps), three complete modulation cycles were presented. Targets and interferers had a level of 50 dB SPL.

All stimuli were generated digitally, played out by a 16-bit digital-to-analog converter (CED 1401 plus) at a sampling rate of 20 kHz, and passed through an anti-aliasing filter (Kemo 21C30) with a cutoff frequency of 8.6 kHz (slope of 100 dB/oct). Stimuli were presented monaurally, using Sennheiser HD250 headphones. Subjects were seated individually in an IAC double-walled sound attenuating booth.

B. Procedure

A two-interval two-alternative forced-choice (2I-2AFC) task was used. The interferer was presented in both intervals in each trial. The subjects were required to indicate the interval with the modulated target as opposed to the unmodulated target. Feedback was provided following each response. The modulation depth (expressed as a percentage of f_c) of the target in the signal interval was adjusted using a two-down one-up adaptive tracking rule (Levitt, 1971), tracking 71% correct responses. At the beginning of each threshold measurement the target was modulated with an FM depth of 10%, which made the task easy in all conditions. Initially, the FM depth was divided or multiplied by a factor of $\sqrt{2}$. Following four reversals, the factor was reduced to $\sqrt[4]{2}$ and eight further reversals were obtained. The threshold estimate was defined as the geometric mean of the FM depths at the last eight reversals. The time interval between presentations of the two target tones in a trial was fixed at 600 ms in all conditions (see Fig. 1). Each interval was marked by a light.

The total duration of a single session was about 2 h, including rest times. At least 12 threshold estimates were obtained for each condition and subject,[1] in a quasirandomized order (one threshold was obtained for each condition in turn, before additional measurements were obtained in any other condition). The order of the conditions was counterbalanced over subjects. Previous work has shown that performance in MDI tasks can improve considerably with practice (Hall and Grose, 1991; Moore *et al.*, 1991; Bacon and Konrad, 1993). Since we were concerned that practice might have differential effects in different conditions, we calculated (i) the geometric mean of the first three thresholds for each condition and subject, and (ii) the geometric mean of the last three thresholds for each condition and subject. To familiarize subjects with the procedure and equipment, they participated in one threshold track in condition none and synchronous FM, respectively, before proper data collection was started.

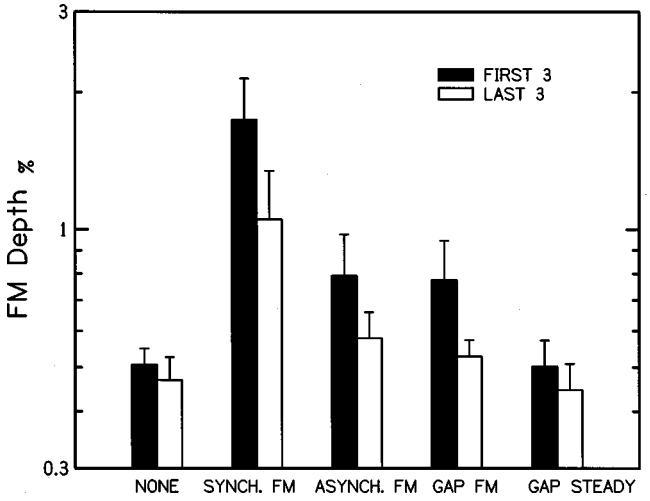


FIG. 2. The geometric mean FM detection thresholds, and the associated geometric standard errors (across subjects). The five groups of bars correspond to the five conditions. The solid and empty bars show the geometric means over the first three threshold estimates and the last three threshold estimates, respectively.

C. Subjects

Seven subjects, one of whom was the first author, participated in all five conditions. Subjects ranged in age from 17 to 38 years, and their quiet thresholds at octave frequencies between 500 and 4000 Hz were within 15 dB of the 1969 ANSI standard.

D. Results and discussion

Figure 2 shows the results averaged over all subjects. The ordinate indicates FM detection thresholds expressed as percentage of the f_c. There are five groups of bars corresponding to the five conditions. Each group consists of two bars. The solid and empty bars show the geometric means of the first three threshold estimates and the last three threshold estimates, respectively. Thresholds were highest when the FM interferer was presented synchronously with the target (synchronous FM). They were lowest when the target was presented alone (none) and when steady tones were presented before and after the target (gap steady). Thresholds

TABLE I. *Post hoc* pairwise comparisons between log-transformed thresholds measured in experiment 1. The significance levels are based on Fisher's least significant difference procedure (see, e.g., Howell, 1997).

Contrast	p
None—asynchronous FM	$\leqslant 0.01$
None—gap FM	$\leqslant 0.01$
None—gap steady	>0.05
Asynchronous FM—gap FM	>0.05

were intermediate when the FM interferer was presented asynchronously but simultaneously with the target (asynchronous FM). When the FM interferer was presented before and after the target but not simultaneously with the target (gap FM), thresholds were similar to those observed for asynchronous FM. Practice effects were observed in all conditions. Thus FMDI occurs when the interferer is not presented at the same time as the target.

To examine the statistical significance of differences between thresholds in the conditions of interest, a two-way ANOVA[2] (Practice×Condition) was calculated which excluded condition synchronous FM. This was done to avoid getting a significant main effect of condition based on the obvious and extensively replicated FMDI for synchronous interferers. The repeated measures ANOVA (Practice[2]×Condition[4]) showed a highly significant main effect of condition [$F(3,18)=9.84$, $p<0.001$],[3] and a significant main effect of practice [$F(1,6)=10.1$, $p<0.05$], but no significant interaction between practice and condition. Moreover, *post hoc* contrasts showed that thresholds in condition asynchronous FM were not significantly different from those in condition gap FM. Thresholds in both conditions were, however, significantly higher than those in conditions gap steady and none (see Table I for details). Note, that the observed effects of condition were unchanged if all 12 threshold estimates were included in the data analysis.

Thus FMDI was observed for interferers which started before and stopped after the target. This indicates that the perceptual segregation of target and interferer due to the onset and offset asynchrony did not completely abolish the interference. This finding is qualitatively similar to the AMDI observed for asynchronous amplitude modulated interferers (Hall and Grose, 1991; Moore and Jorasz, 1992; Moore and Shailer, 1992; Mendoza *et al.*, 1995). It suggests that FMDI like AMDI, is probably comprised of more than one component. The reduction found with onset and offset asynchrony probably represents that portion of FMDI attributable to auditory grouping by onset/offset synchrony. Interestingly and more importantly, the simultaneous part of the FM interferer was not necessary to produce FMDI. The gapped FM interferer produced about the same amount of impairment as the asynchronous continuous FM interferer. This finding contradicts Yost and Sheft's (1989) model of MDI, which suggests that MDI is the consequence of simultaneously present modulated carriers interacting in specialized channels for detecting temporal modulation. The fact that FMDI does not occur only for stimuli which overlap in time with the target, indicates that FMDI cannot be explained completely in terms of fusion, i.e., a noneffective segregation of *concurrent* sounds.

Before thinking about which aspects of *sequential* sound segregation might play a role in FMDI, there are two important points to consider. In this experiment, all stimuli had relatively short onset and offset ramps of 10 ms. Also, in the gapped conditions, the target started when the first part of the interferer stopped, and the second part of the interferer started when the target stopped. It is conceivable that, under these conditions, the interferer and target were not clearly perceived as separate auditory objects, even though their carrier frequencies differed substantially. Subjectively, the FM seemed to ''connect'' interferer and target to a certain degree. Since the FM of the target continued in phase with the interferer, this might lead to a kind of continuous FM perception.

Additionally, there is the possibility that there is ringing of filters in a hypothetical modulation filter bank (Houtgast, 1989; Yost *et al.*, 1989; Dau *et al.*, 1997a, 1997b). This filter bank has usually been discussed in the literature as a bank of filters selective for *amplitude* modulation. However, within individual auditory filters FM is translated into AM (Zwicker, 1956). Also, Moore *et al.* (1991) showed that AM interferers impair detection of a change in the depth of FM targets, and vice versa. From this they concluded that there are no separate channels for AM and FM. Therefore, if a modulation filter bank exists, then ringing in its filters would quite likely occur for a certain time after presenting a FM tone. Thus one might argue that thresholds for FM detection in the presence of *physically* nonsimultaneous FM interferers increased because of some kind of *internal* persistence at the level of the modulation filters.

These two points are not mutually exclusive. Actually, ringing in the modulation filters might be one reason for a more connected, instead of independent perception of target and interferers. In any case, one should remember that segregation is not all or none (Gockel *et al.*, 1999). Therefore, introducing a silent gap between target and interferer in the current experiment is likely to enhance target and interferer being perceived as separate auditory objects.

The second experiment investigated whether FMDI for nonsimultaneous interferers would persist, even when the interferer and target were separated by a silent interval. Several durations of the silent interval were tested. The durations were chosen in such a way that they would allow decay of any ringing in the hypothetical modulation filterbank, and enhance segregation of the stimuli.

III. EXPERIMENT 2: VARIATION OF SILENT GAP DURATION

A. Stimuli

The basic stimuli were the same as in experiment 1. However, instead of having just one gapped FM interferer condition (with an interval of 200 ms between the parts of the interferer which precede and follow the target, as was the case in experiment 1), the duration of the silent gap was varied between conditions. The rationale was that if the time between offset of the precursor part of the FM interferer and

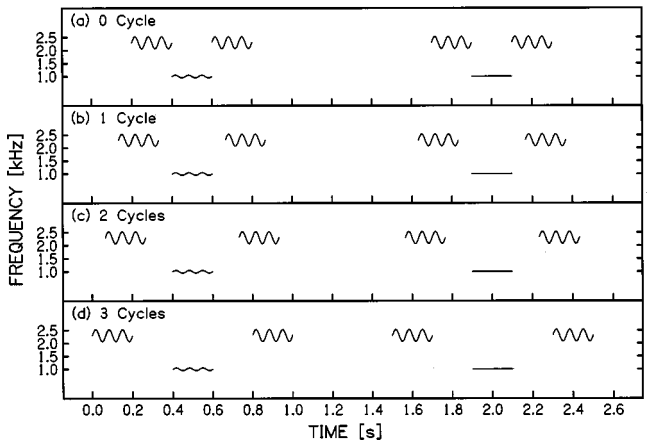


FIG. 3. Schematic spectrograms of stimuli presented over the course of one 2AFC trial in the different conditions: (a) 0 cycle, (b) 1 cycle, (c) 2 cycles, (d) 3 cycles. Note that for illustrative purposes only (i) an extreme FM depth has been chosen for the target, and (ii) the signal interval is always the first interval.

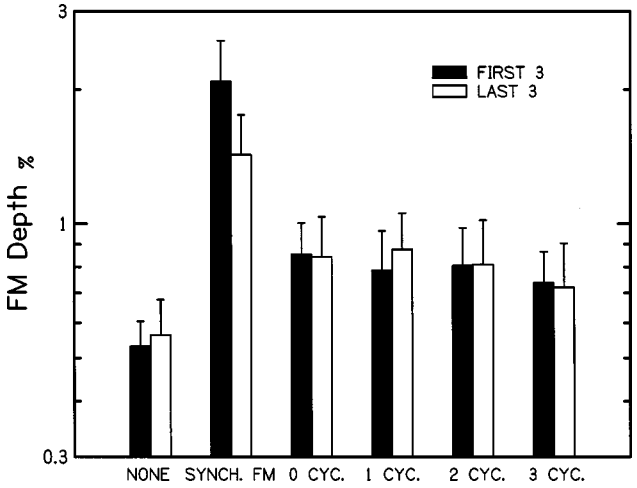


FIG. 4. The geometric mean FM detection thresholds, and the associated geometric standard errors (across subjects). The six groups of bars correspond to the six conditions. The solid and empty bars show the geometric means over the first three threshold estimates and the last three threshold estimates, respectively.

onset of the target, and similarly the time between offset of the target and onset of the postcursor part of the FM interferer was made long enough, then the stimuli should be perceived clearly as separate auditory events. If FMDI is observed in such a situation, then this could not be explained by subjects not being able to clearly identify the beginning and end of the target tone. Another reason for introducing silent gaps of various durations between the stimuli was that this would make it possible to assess the contribution of ringing in a hypothetical modulation filter bank to the FMDI produced by nonsimultaneous interferers.

Four values of the silent gap duration on either side of the target were used. They were chosen to correspond to a whole number of modulation cycles (see Fig. 3): (1) 0 ms (condition 0 cycle), (2) 66.67 ms (condition 1 cycle), (3) 133.33 ms (condition 2 cycles), and (4) 200 ms (condition 3 cycles). Whole periods of modulation cycles were chosen, to keep constant the phase relationship between the ringing in the modulation filter centered on 15 Hz and the response of the same filter to the next stimulus. This prevents the occurrence of dips of unintended variable duration in the output of the modulation filter; see Shailer and Moore (1987) for a description of such effects in the carrier frequency domain. Condition 0 cycle is similar to the gap FM interferer condition in experiment 1. In addition to these four gap interferer conditions, performance was again measured for targets presented alone (condition ''none''), and for targets presented simultaneously and synchronously with FM interferers (condition synchronous FM). For all conditions, the time between presentations of the target tones in the two intervals was 1300 ms (see Fig. 3). Each interval was marked by a light.

B. Subjects and procedure

Seven new subjects participated in all six conditions. Subjects ranged in age from 18 to 39 years, and their quiet thresholds at octave frequencies between 500 and 4000 Hz were within 15 dB of the 1969 ANSI standard. At least 10 threshold estimates were obtained for each condition and subject[4] in a quasirandomized order (one threshold was obtained for each condition in turn, before additional measurements were obtained in any other condition). Apart from this, the procedure was the same as for experiment 1. The order of the conditions was counterbalanced over subjects. To check on practice effects, the geometric mean of the first three thresholds for each condition and subject, and the geometric mean of the last three thresholds for each condition and subject were calculated.

C. Results and discussion

Figure 4 shows the results averaged over all subjects. There are six groups of bars corresponding to the six conditions. The solid and empty bars show the geometric means of the first three threshold estimates and the last three threshold estimates, respectively. FM detection thresholds were highest for the synchronous interferer. They were lowest when the target was presented alone (none). Thresholds were intermediate for the various gapped FM interferers. Practice effects were quite varied.

To examine the statistical significance of differences between thresholds in the conditions of interest, a two-way ANOVA (Practice×Condition) was calculated which excluded condition synchronous FM, for the same reason as described for experiment 1. The repeated measures ANOVA (Practice[2]×Condition[5]) showed a highly significant main effect of condition [$F(4,24)=15.54$, $p<0.001$], but no significant effect of practice. *Post hoc* contrasts indicated that thresholds in all gap interferer conditions (for all gap durations) differed significantly from thresholds in condition none ($p\leqslant 0.005$ in all comparisons). Thus FMDI was observed for all durations of the silent gaps between FM interferers and target. Note, that the observed effects were unchanged if all ten threshold estimates were included in the data analysis. There was a slight tendency for the mean thresholds to decrease with increase in duration of the silent gap (see Table II), but this was not significant.

As mentioned above (see Sec. II D), FM is translated into AM within individual auditory filters. Thus physically

TABLE II. Mean thresholds measured in experiment 2 specified as peak deviation from f_c expressed as percentage of the f_c.

Condition	Mean
None	0.547
3 cycles	0.700
2 cycles	0.766
1 cycle	0.789
0 cycle	0.811

nonsimultaneous FM interferers theoretically could lead to FMDI due to some kind of internal persistence at the level of the filters in an amplitude modulation filter bank. In order to assess the possible effects of ringing in an amplitude modulation filter bank it was necessary to calculate the effective AM depth of the interferer. To do this we calculated the excitation patterns (Glasberg and Moore, 1990) for sinusoids with frequencies of 2530 and of 2070 Hz ($2300 \pm 10\%$, i.e., at the extremes of modulation cycle), at 50 dB SPL. The maximum difference between the two excitation patterns occurs on the low-frequency side (around 1770 Hz) and is about 33 dB. Thus the maximum amplitude modulation within a single auditory channel resulting from our FM interferer corresponds approximately to a ratio of $45(10^{33/20})$ between maximum amplitude and minimum amplitude. This corresponds to a modulation index, m, of 0.956, which, in terms of $20 \log m$, equals -0.39 dB. The absolute threshold for detection of amplitude modulation reported in the literature is around -25 dB (expressed as $20 \log m$). Thus the ringing in the modulation filter centered on 15 Hz (remember the f_m rate equaled 15 Hz) has to decay by about 25 dB for the amplitude modulation to be below detection threshold.

We then estimated how much the ringing of an amplitude modulation filter centered at 15 Hz would have decayed at a certain time after the end of input (after presentation of the interferer). The modulation filters are very broadly tuned, and Dau *et al.* (1997a) assume a second-order filter with a Q value of two for center frequencies above 10 Hz, where Q is the center frequency divided by the bandwidth (BW). The time constant (the time by which the ringing of a filter has decayed by a factor of $1/e$) of a second order bandpass filter is $1/(\pi \mathrm{BW})$ [see, e.g., Morse (1983); Hartmann (1997)]. For a Q value of two, after two cycles the ringing has decayed by 27 dB and after three cycles it has decayed by 40 dB. This means that for conditions 2 cycles and 3 cycles, the ringing of the modulation filter would have decayed below detection threshold. Thus there should be no perceptible interference in a modulation filterbank. Therefore, the interference found for the gapped interferer cannot be explained in terms of ringing in a modulation filter bank.

Furthermore, the silent gap of 200 ms between interferers and target subjectively led to a clear distinction between offset of one stimulus and onset of a new stimulus. So, overall, these results demonstrate that asynchronous, nonsimultaneous FM interferers can impair detection of FM of a target tone even if interferer and target are clearly separated not only in frequency but also in time, and therefore are perceived as distinct auditory objects. This finding contradicts Yost and Sheft's (1989) model of MDI, which suggests that MDI is the consequence of simultaneously present modulated carriers interacting in specialized channels for detecting temporal modulation.

IV. OVERALL DISCUSSION

The first experiment showed that the amount of FMDI decreased when an onset and offset asynchrony between interferer and target was introduced. This reduction in FMDI for asynchronous stimuli is qualitatively similar to findings reported earlier for AMDI (Hall and Grose, 1991; Moore and Jorasz, 1992; Moore and Shailer, 1992; Mendoze *et al.*, 1995). Remarkably, the amount of FMDI produced by asynchronous interferers was the same, whether or not some part of the interferer was presented simultaneously with the target. In contrast, a steady asynchronous and nonsimultaneous interferer did not impair performance. The second experiment demonstrated that the FMDI produced by the asynchronous gapped interferers did not depend on some sort of interaction due to ringing in modulation filters. It also showed that FMDI persisted, even when interferers and targets were clearly separated in frequency *and time* and were perceived as separate auditory objects.

A finding which might be related was recently reported by Cusack and Carylon (2000). They presented listeners sequentially with three 500-ms sounds. In one condition, two of the sounds were steady sinusoids and one (the second or third) was a FM tone. In the other condition, two of the sounds were FM tones and one (the second or third) was a steady sinusoid. All tones had the same carrier frequency and were modulated (if they were modulated) with the same FM depth (resulting in 71% correct responses in a one interval task, as determined individually before the experiment proper). The task of the subjects was to identify the ''odd-man-out.'' They found that performance was better in identifying the single FM tone than in identifying the single steady tone (difference in d' of about 1). There is some similarity between our stimuli and the ones used in their study. In condition gap steady in experiment 1, a sequence of three stimuli—steady tone, FM tone, steady tone—was presented in the signal interval. In the standard interval all stimuli were steady. No FMDI was observed in this condition. On the other hand, in condition gap FM, a sequence of FM tone, steady tone, FM tone, was presented in the standard interval. In the signal interval all stimuli were FM tones. This condition led to FMDI. So, in the present experiment, we also found performance to be better in the condition where subjects had to identify the observation interval which contained an odd-man-out that was frequency modulated. Note that, from this point of view, in condition gap FM subjects would have to choose the interval which did *not* contain the odd-man-out sequence. If subjects used this strategy, it would still have led to performance being worse when the odd-man-out was a steady tone (condition gap FM) than when the odd-man-out was a FM tone (condition gap steady). This is analogous to the findings of Cusack and Carlyon (2000).

Cusack and Carlyon discussed their results in terms of a perceptual asymmetry of FM tones and steady tones. They assumed FM tones to ''pop out'' among other *sequentially*

presented steady tones because of some additional feature (the FM) which makes them more salient. This interpretation is supported by earlier findings by McAdams (1989). He investigated the perceived prominence of individual target vowels *presented simultaneously* with other vowels. He found that FM did increase the prominence of the target vowel compared to when it was not frequency modulated. Interestingly, this was true independent of whether or not the other vowels were also modulated. The relative phase of modulation of the target vowel and the nontarget vowels had no effect. Thus for concurrently presented complex tones, there is evidence for an increase in salience of those parts which are frequency modulated. In line with this, Summerfield and Culling (1992) found that segregation of concurrent vowels was helped by modulating the frequency components of one vowel when the other was not modulated (provided both vowels were composed of components placed randomly in frequency). It is quite conceivable that the asymmetry in salience between FM and steady tones also leads to performance differences in discrimination experiments with *sequentially* presented stimuli, similar to the ones used by Cusack and Carlyon.

Such an asymmetry in salience between FM and steady tones could be the reason why, in the present study, performance in condition gap FM was worse than in condition gap steady. In the gap FM condition, the FM interferer would be very salient, i.e., would ''pop out,'' and draw attention away from the target modulation. In the gap steady condition, only the target was modulated. It therefore would pop out among the steady tones, leading to improved performance. An asymmetry in salience might also explain why introducing the asynchrony between interferer and target only reduced the amount of FMDI, but did not abolish it. The asynchrony would enhance perceptual segregation and reduce the amount of FMDI when compared to the synchronous condition. However, the difference in salience between the strongly modulated interferer and the near-threshold modulated target would prevent the complete abolition of FMDI. Following the same line of argument, asymmetry in salience between the strongly modulated interferer and the target with a much less degree of modulation could contribute to the FMDI observed with synchronous interferers.

Theoretically, one could argue that performance in the conditions with gapped FM interferers might be impaired due to adaptation effects caused by the interferer in channels tuned to FM (Kay and Mathews, 1972; Gardner and Wilson, 1979; Tansley and Suffield, 1983; Wakefield and Viemeister, 1984). However, this is quite unlikely, for several reasons: (i) Kay and Matthews (1972) measured FM detection thresholds as a function of the duration of an FM adaptor which had a f_c identical to that of the target. The shortest adaptor duration used was 5 s. For a 200 ms-adaptor (equal to the duration of the interferer in the present study), the curve they fitted to their data would predict an increase in thresholds by a factor of about 1.2 which is less than that actually observed in our study (factor of 1.48 for the 0 cycle condition). (ii) Kay and Matthews (1972) also measured FM detection thresholds as a function of the separation of the f_cs of the adaptor and the target. The ''filter'' function they fitted to their data (for a 12-s adaptor) showed a maximum increase in thresholds by a factor of 2.5 when adaptor and target had identical f_c. For an adaptor with f_c 1.3 octaves higher than that of the target, their curve predicts no increase in threshold due to adaptation. Since in the present study the f_c of the interferer was 1.3 octaves higher than that of the target, there should have been no threshold elevation due to adaptation (even for much longer interferer durations). (iii) For short adaptors (duration below about 6 to 12 s) thresholds increase with increase in duration of the adaptor (Kay and Matthews, 1972). If adaptation were the cause of the increased thresholds in the conditions with gapped FM interferers, one would expect thresholds to be higher in the asynchronous than in the gapped condition, due to the increased duration of the interferer in the former condition. This was not the case.

Since FM is translated into AM at the output of individual auditory filters, one could theoretically also argue that adaptation in filters tuned to AM (caused by the FM interferer) was the reason for the impaired performance in the gapped FM conditions. However, this is quite unlikely for the following reasons: (i) Even though our excitation pattern calculations showed that the interferer would produce sizable AM on the low-frequency side of its excitation pattern, this occurred at a frequency separated by about one octave from the detectable AM on the low-frequency side of the target's excitation pattern (calculated for a target with 0.5% FM depth, corresponding to the average threshold in condition none). Richards *et al.* (1997) measured the increase in AM detection thresholds after prolonged exposure to AM adaptors as a function of the separation of the f_cs of the target and the adaptor. For a frequency separation of about one octave (f_c of target below f_c of adaptor), they observed an increase in thresholds which was smaller than the size of our effect, defined as the increase in the value of m for the target at 0.5 and 0.8% FM depth, respectively. (ii) For AM stimuli, changes in the excitation pattern are largest on the high-frequency side (Glasberg and Moore, 1990). Thus at presentation levels of 65 dB SPL, as used by Richards *et al.*, an adaptor with an f_c one octave above that of the target would lead to significant modulation in excitation at the output of the auditory filters providing the most reliable cues for detection of AM of the target. In contrast, for FM stimuli, changes in the excitation pattern are largest on the low-frequency side (Zwicker, 1956). Therefore, for a given frequency separation of the target and the adaptor (with the adaptor higher in frequency than the target) the effect of the adaptor should be markedly smaller for FM-induced AM than for AM. Overall, we conclude that our results cannot be explained in terms of AM adaptation.

V. SUMMARY AND CONCLUSIONS

The main findings of this study are:

(1) The FMDI (increase in threshold relative to that obtained for no interferer) observed for the synchronous interferer condition was reduced but not eliminated by introducing an onset and offset asynchrony between interferer and target. This extends earlier observations of MDI from AM to FM.

(2) The amount of FMDI for asynchronously gated interferers did not depend on whether or not they continued through the presentation of the target. No FMDI was observed for an asynchronous gapped steady interferer.
(3) Increasing the duration of the silent gap (centered on the target) between the two parts of the interferer did not eliminate the FMDI.

We conclude that asynchronous, nonsimultaneous FM interferers can impair detection of FM of a target tone even if the interferer and target are clearly separated in frequency and time, and therefore are perceived as distinct auditory objects. The results are consistent with the idea that the higher salience of FM tones compared to steady tones leads to impaired performance in a discrimination task even for sequentially presented stimuli.

ACKNOWLEDGMENTS

The authors wish to thank Helga Sukowski for collecting part of the data in experiment 1, and Brian Glasberg and Brian Moore for providing us with the program for the calculation of excitation patterns. We also thank Brian Moore, Rob Hukin, and two anonymous reviewers for helpful comments on a previous version of this manuscript.

[1]All subjects were run for the same amount of time (number of sessions). This resulted in 12 threshold estimates per condition for one subject, 15 for five subjects, and 16 for one subject.
[2]Repeated measurements ANOVAs were run on the logarithms of the geometric mean thresholds of each subject.
[3]Throughout the paper, if appropriate, the Huynh–Feldt correction was applied to the degrees of freedom (see, e.g., Howell, 1997).
[4]All subjects were run for the same amount of time (number of sessions). This resulted in 10 threshold estimates per condition for one subject, 11 for another subject, 12 for four subjects, and 14 for one subject.

Bacon, S. P., and Konrad, D. L. (**1993**). ''Modulation detection interference under conditions favoring within- or across-channel processing,'' J. Acoust. Soc. Am. **93**, 1012–1022.
Bacon, S. P., and Moore, B. C. J. (**1993**). ''Modulation detection interference: Some spectral effects,'' J. Acoust. Soc. Am. **93**, 3442–3453.
Bacon, S. P., Moore, B. C. J., Shailer, M. J., and Jorasz, U. (**1995**). ''Effects of combining maskers in modulation detection interference,'' J. Acoust. Soc. Am. **97**, 1847–1853.
Bacon, S. P., and Opie, J. M. (**1994**). ''Monotic and dichotic modulation detection interference in practiced and unpracticed subjects,'' J. Acoust. Soc. Am. **95**, 2637–2641.
Carlyon, R. P. (**1994**). ''Further evidence against an across-frequency mechanism specific to the detection of frequency modulation (FM) incoherence between resolved frequency components,'' J. Acoust. Soc. Am. **95**, 949–961.
Cusack, R., and Carlyon, R. P. (**2000**). ''Auditory pop-out: Perceptual asymmetries in sequences of sounds,'' Br. J. Audiol. **34.2**, 112.
Darwin, C. J., and Ciocca, V. (**1992**). ''Grouping in pitch perception: Effects of onset asynchrony and ear of presentation of a mistuned component,'' J. Acoust. Soc. Am. **91**, 3381–3390.
Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997a**). ''Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers,'' J. Acoust. Soc. Am. **102**, 2892–2905.
Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997b**). ''Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration,'' J. Acoust. Soc. Am. **102**, 2906–2919.
Gardner, R. B., and Wilson, J. P. (**1979**). ''Evidence for direction-specific channels in the processing of frequency modulation,'' J. Acoust. Soc. Am. **66**, 704–709.
Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hear. Res. **47**, 103–138.
Gockel, H. (**2000**). ''Perceptual grouping and pitch perception,'' in *Results of the 8th Oldenburg Symposium on Psychological Acoustics*, edited by A. Schick, M. Meis, and C. Reckhardt (BIS, Oldenburg, Bad Zwischenahn, Germany).
Gockel, H., Carlyon, R. P., and Micheyl, C. (**1999**). ''Context dependence of fundamental-frequency discrimination: Lateralized temporal fringes,'' J. Acoust. Soc. Am. **106**, 3553–3563.
Hall, J. W., and Grose, J. H. (**1991**). ''Some effects of auditory grouping factors on modulation detection interference (MDI),'' J. Acoust. Soc. Am. **90**, 3028–3035.
Hartmann, W. M. (**1997**). *Signals, Sound, and Sensation* (AIP, Woodbury, New York).
Houtgast, T. (**1989**). ''Frequency selectivity in amplitude-modulation detection,'' J. Acoust. Soc. Am. **85**, 1676–1680.
Howell, D. C. (**1997**). *Statistical Methods for Psychology*, 4th ed. (Duxbury, Belmont, CA).
Kay, R. H., and Mathews, D. R. (**1972**). ''On the existence in human auditory pathways of channels selectively tuned to the modulation present in frequency-modulated tones,'' J. Physiol. (London) **225**, 657–677.
Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.
Lyzenga, J., and Carlyon, R. P. (**1999**). ''Center frequency modulation detection for harmonic complexes resembling vowel formants and its interference by off-frequency maskers,'' J. Acoust. Soc. Am. **105**, 2792–2806.
McAdams, S. (**1989**). ''Segregation of concurrent sounds. I Effects of frequency modulation coherence,'' J. Acoust. Soc. Am. **86**, 2148–2159.
Mendoza, L., Hall, J. W., and Grose, J. H. (**1995**). ''Within- and across-channel processes in modulation detection interference,'' J. Acoust. Soc. Am. **97**, 3072–3079.
Moore, B. C. J., Glasberg, B. R., Gaunt, T., and Child, T. (**1991**). ''Across-channel masking of changes in modulation depth for amplitude- and frequency-modulated signals,'' Q. J. Exp. Psychol. **43A**, 327–347.
Moore, B. C. J., and Jorasz, U. (**1992**). ''Detection of changes in modulation depth of a target sound in the presence of other modulated sounds,'' J. Acoust. Soc. Am. **91**, 1051–1061.
Moore, B. C. J., and Shailer, M. J. (**1992**). ''Modulation discrimination interference and auditory grouping,'' Philos. Trans. R. Soc. London, Ser. B **336**, 339–346.
Morse, P. M. (**1983**). *Vibration and Sound* (American Institute of Physics, Massachusetts).
Richards, V. M., Buss, E., and Tian, L. (**1997**). ''Effects of modulator phase for comodulation masking release and modulation detection interference,'' J. Acoust. Soc. Am. **102**, 468–476.
Shailer, M. J., and Moore, B. C. J. (**1987**). ''Gap detection and the auditory filter: Phase effects using sinusoidal stimuli,'' J. Acoust. Soc. Am. **81**, 1110–1117.
Summerfield, Q., and Culling, J. (**1992**). ''Auditory segregation of competing voices: Absence of effects of FM or AM coherence,'' Philos. Trans. R. Soc. London, Ser. B **336**, 357–366.
Tansley, B. W., and Suffield, J. B. (**1983**). ''Time course of adaptation and recovery of channels selectively sensitive to frequency and amplitude modulation,'' J. Acoust. Soc. Am. **74**, 765–775.
Wakefield, G. H., and Viemeister, N. F. (**1984**). ''Selective adaptation to linear frequency-modulated sweeps: Evidence for direction specific channels?,'' J. Acoust. Soc. Am. **75**, 1588–1592.
Wilson, A. S., Hall, J. W., and Grose, J. H. (**1990**). ''Detection of frequency modulation (FM) in the presence of a second FM tone,'' J. Acoust. Soc. Am. **88**, 1333–1338.
Yost, W. A., and Sheft, S. (**1989**). ''Across-critical-band processing of amplitude-modulated tones,'' J. Acoust. Soc. Am. **85**, 848–857.
Yost, W. A., and Sheft, S. (**1990**). ''A comparison among three measures of cross-spectral processing of amplitude modulation with tonal signals,'' J. Acoust. Soc. Am. **87**, 897–900.
Yost, W. A., Sheft, S., and Opie, J. (**1989**). ''Modulation interference in detection and discrimination of amplitude modulation,'' J. Acoust. Soc. Am. **86**, 2138–2147.
Zwicker, E. (**1956**). ''Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs'' Acustica **6**, 356–381.

Effects of relative phase and frequency spacing on the detection of three-component amplitude modulation

Brian C. J. Moore[a)]
Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, England

Aleksander Sek
Institute of Acoustics, Adam Mickiewicz University, 85 Umultowska, 61-614 Poznan, Poland

(Received 15 February 2000; revised 12 May 2000; accepted 2 August 2000)

These experiments explored the effect of relative modulator phase on the detection of a three-component modulator applied to a 4000-Hz sinusoidal carrier with a level of 70 dB SPL. The central modulator component had a frequency of 50 Hz, and the two other components had frequencies of 50 ± 5, 10, 25, 40, or 45 Hz. Thus, the modulator waveform was always periodic. Each modulator component had the same modulation index, *m*. The relative phases of the components were chosen to give a variety of modulation waveforms differing in the ratio of maximum to minimum value (max–min) and in crest factor. In experiment 1, modulation detection thresholds were measured by varying *m*, using an adaptive two-interval forced-choice procedure. Thresholds were found to be independent of relative modulator phase and of the frequency spacing of the components. In experiment 2, detectability (d') of the modulation was measured for several fixed values of *m*. Detectability was found to be independent of relative modulator phase and of the frequency spacing of the components. The results are not consistent with the idea that modulation detection thresholds are determined by the max–min value or crest factor of the envelope. The results are consistent with a model which assumes that the stimuli are subjected to a nonlinearity, and thresholds are determined by the root-mean-square value (or the mean square value) of the ac component of the envelope, following this nonlinearity. The nonlinearity may partly reflect compression on the basilar membrane, but other nonlinearities may be involved. This model can also explain some aspects of earlier results on the sensitivity to relative modulator phase [E. A. Strickland and N. F. Viemeister, J. Acoust. Soc. Am. **99**, 3638–3646 (1996)]. © *2000 Acoustical Society of America.* [S0001-4966(00)01411-9]

PACS numbers: 43.66.Mk, 43.66.Nm, 43.66.Ba [DWG]

I. INTRODUCTION

The perception of amplitude modulation (AM) has often been modeled by a sequence of stages. The first stage is a bank of bandpass filters (the auditory filters), although some models assume just a single, broad ''predection filter'' (Viemeister, 1979). Each filter is followed by a nonlinear device, such as a rectifier, and a ''smoothing'' device. Usually, this smoothing device is though to reflect a relatively central process. The output of the smoothing device is fed to a decision device. Models of this general type have been proposed by Viemeister (1979), Festen and Plomp (1981), Moore *et al.* (1988), Plack and Moore (1990), and Oxenham and Moore (1994). The smoothing device has been implemented as a low-pass filter (Viemeister, 1979; Strickland and Viemeister, 1996) or by a sliding temporal integrator (temporal window) (Moore *et al.*, 1988; Plack and Moore, 1990; Oxenham and Moore, 1994). The decision mechanism has been based on various properties of the output of the smoothing device, including the root-mean-square (rms) value (Viemeister, 1979), the crest factor (the ratio of the maximum value to the rms value) (Hartmann and Pumplin, 1988; Strickland and Viemeister, 1996; Lorenzi *et al.*, 1999) and the ratio of the maximum value to the minimum value (max–min) (Forrest and Green, 1987; Strickland and Viemeister, 1996).

One goal of the present study was to assess the adequacy of decision statistics based on the crest factor or max–min value in accounting for the detection of multicomponent modulators. We used three-component modulators whose relative phase was chosen to give modulator waveforms which, for a given rms value, differed in max–min value and crest factor (see below for details). If modulation detection thresholds depend on the max–min value or crest factor, modulation detection thresholds, expressed as rms values, should vary with the relative phase of the modulator components.

The effect of relative modulator phase has been assessed in only a few previous studies, and these mostly involved ''modulation masking'' (detection of signal modulation in the presence of masker modulation) rather than absolute thresholds for detecting complex modulation. Bacon and Grantham (1989) measured thresholds for detecting sinusoidal AM of a white-noise carrier in the presence of a second sinusoidal AM, acting as a modulation masker. They found no effect of relative phase using masker and signal modulation frequencies of 6 and 4 Hz, respectively. Strick-

[a)]Author to whom correspondence should be addressed. Electronic mail: bcjm@cus.cam.ac.uk

land and Viemeister (1996) also examined modulation masking using a noise carrier and a sinusoidal masker and a sinusoidal signal. When the two sinusoids had a frequency ratio of 2:1, thresholds were influenced by the relative phase of the masker and signal. For example, when the masker frequency was twice the signal frequency and the masker frequency was relatively low (32 or 8 Hz), thresholds tended to be lowest (best) when the starting phase of the signal relative to the masker was 135° or 225°. Strickland and Viemeister were able to account for the general form of the phase effect (but not its absolute magnitude) within the framework of the model of Viemeister (1979), using the max−min value as a decision statistic. They also pointed out that, consistent with the experimental results, this statistic leads to the prediction of no phase effects for the conditions tested by Bacon and Grantham (1989).

Lorenzi *et al.* (1999) studied the effect of relative modulator phase for a two-component modulator when both components were clearly detectable. A white-noise carrier was modulated with the sum of two sinusoids with frequency ratio 2:1 (100, 200 and 150, 300) or 3:1 (100, 300). The task was to distinguish two stimuli in which the two sinusoidal modulators had either a fixed phase or a drifting phase. Performance was good for the 2:1 ratio, but at chance for the 3:1 ratio. Simulations using an envelope-detector model like that of Viemeister (1979) and Strickland and Viemeister (1996) showed that a max−min statistic predicted above-chance performance for the ratio of 3:1. In other words, the max−min statistic predicted a sensitivity to relative phase when in fact there was none. The data could be accounted for using as a statistic either the crest factor or the skewness of the envelope.

The present paper was intended to contribute further to knowledge of the sensitivity of the auditory system to relative modulator phase, and to evaluate the adequacy of the max−min and crest factor statistics in predicting absolute thresholds for the detection of modulation. We measured thresholds for detecting amplitude modulation (experiment 1) or the detectability of amplitude modulation (experiment 2) as a function of relative modulator phase, using a three-component modulator. By manipulating the relative phase of the components we could create a variety of modulator waveforms, all with the same root-mean-square (rms) value but with differing max−min ratios and crest factors.

A second goal of our study was to test the adequacy of a model of modulation perception proposed by Dau and coworkers (Dau, 1996; Dau *et al.*, 1997a, 1997b). This model is based on the assumption that modulation perception depends on a second stage of filtering in the auditory system. It is assumed that the envelopes of the outputs of the auditory filters are fed to a second array of overlapping bandpass filters tuned to different envelope modulation rates. This set of filters is often called a ''modulation filter bank'' (MFB). The concept of the MFB implies that the auditory system performs a limited resolution spectral analysis of the envelope at the output of each auditory filter. Psychoacoustical evidence supporting the concept of the MFB has come from experiments on modulation masking (Houtgast, 1989; Bacon and Grantham, 1989; Strickland and Viemeister, 1996; Dau *et al.*, 1997a), and from experiments on AM detection with various carrier bandwidths and noise types (Dau *et al.*, 1997a, 1997b, 1999). These have revealed selectivity in the modulation domain.

The model proposed by Dau *et al.* (1997a) is based on the assumption that, for center frequencies above 10 Hz, the modulation filters have a Q value (center frequency divided by bandwidth) of 2. Furthermore, for center frequencies above 10 Hz, only the envelope at the output of the modulation filter is assumed to be relevant. Therefore, if a modulator is composed of two or more sinusoids with widely spaced frequencies, so that the sinusoids are effectively resolved by the MFB, the perception of the modulation should not be affected by the relative phase of the sinusoids.

At first sight, the phase effects found by Strickland and Viemeister (1996) appear inconsistent with Dau's model. However, within the framework of Dau's model, modulation masking only occurs if the masker produces sufficient output from the (broadly tuned) modulation filter used to detect the signal. Therefore, the phase effect observed for the ratio of 2:1 would have to reflect an interaction of the masker and signal within one modulation filter; the results do not provide a test of the sensitivity to relative modulator phase across different modulation filters.

In an attempt to assess the adequacy of the MFB model, we manipulated the frequency spacing of the components. If there is an MFB, then for small spacings, all three components should interact within a single modulation filter, whereas for large spacings such interactions should not occur, especially as all our modulator components had relatively small modulation depths. We anticipated that, for small spacings, we might observe effects of relative phase related to differences in max−min ratio or crest factor, but that for large spacings the phase effects might disappear.

In contrast to the experiments described above, we used a 4-kHz sinusoidal carrier rather than a noise carrier. The sinusoidal carrier has the advantage that the results are not influenced by the intrinsic fluctuations that play a role for a noise carrier (Dau *et al.*, 1997a, 1997b). However, when a sinusoidal carrier is used, there is a danger that spectral changes associated with the modulation might be detectable. To reduce the likelihood that spectral changes would be detected, the highest modulation frequency was limited to 95 Hz. Since the auditory filter bandwidth at 4 kHz is about 456 Hz (Glasberg and Moore, 1990), this made it very unlikely that any spectral sideband could be resolved; see Dau (1996).

II. EXPERIMENT 1: MODULATION DETECTION THRESHOLDS

A. Stimuli

We used a three-component modulator in which each component was of equal amplitude. We did this because we wanted to explore the effects of modulator phase at depths close to the detection threshold, and for a 4-kHz carrier detection thresholds are roughly independent of modulation rate for rates up to about 100 Hz (Dau, 1996). We started by exploring the range of modulator waveforms that could be obtained using three equal-amplitude components with a har-

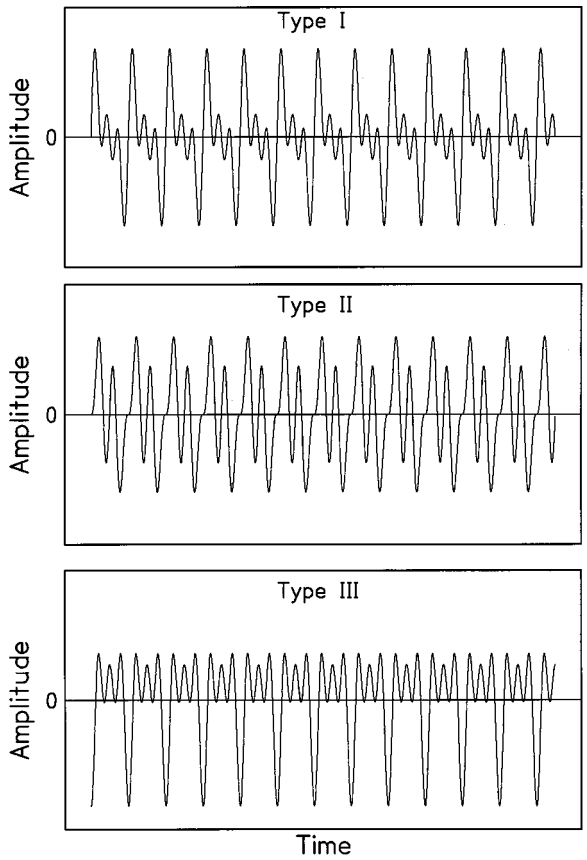


FIG. 1. Schematic illustration of the modulator waveforms for the three modulator types; only the ac components of the modulators are shown. For each panel, the modulator consisted of three equal-amplitude components with frequencies of 25, 50, and 75 Hz. The starting phases of the components were, 0, 0, and 0° for type I, 0, 0, and 180° for type II, and 270, 270, and 270° for type III.

monic relationship. We chose to use three different sets of starting phases that gave modulator waveforms differing in degree of symmetry, crest factor, and max–min value; these will be called types I, II, and III. Examples of the modulator waveforms are illustrated in Fig. 1. In each panel, the modulator components are the same (25, 50, and 75 Hz) and the rms modulation depth is the same; the relative phases of the components vary across panels. Only the ac component of each modulator is shown.

For type I (top panel), the starting phase was 0° (sine phase) for all modulator components. For type II, the highest component was phase-shifted by 180° (the other two components starting at 0°). This gave a less ''peaky'' modulator waveform. For type III, all components had a starting phase of 270°. This gave a more asymmetric modulator waveform. The center component of the modulator always had a frequency of 50 Hz. The spacing of the components was 5, 10, 25, 40, or 45 Hz, so that the three component frequencies were always harmonically related and the compound waveform was periodic. Table I gives the values of the maxima and minima in the modulator waveform, relative to the rms value, for each modulator type and for each frequency spacing of the components. Generally, for a given rms value, the type I and type III modulators had greater max–min ratios than the type II modulator.

In each observation interval, two stimuli were presented, one modulated and one unmodulated; the order of the two was random. Each stimulus had an overall duration of 1000 ms, including 20-ms raised-cosine rise/fall times. When present, the modulation occurred through the whole duration of the stimulus. The interstimulus interval was 400 ms. The level of the carrier was 70 dB SPL.

Stimuli were generated using a Tucker-Davis array processor (TDT-AP2) in a host PC, and a 16-bit digital-to-analog converter (TDT-DD1) operating at a 50-kHz sampling rate. They were attenuated (TDT-PA4) and sent through an output amplifier (TDT-HB6) to a Sennheiser HD414 earphone.

B. Procedure

Thresholds were measured using an adaptive two-interval, forced-choice procedure, with a two-down, one-up stepping rule that estimates the 70.7%-correct point on the psychometric function. The carrier was unmodulated in one interval of a trial (selected at random) and modulated in the other. The task of the subject was to indicate, by pressing one of two buttons, the interval containing the modulation. Feedback was provided by lights following each response. At the start of a run, the modulation depth was chosen to be well above the threshold value. Following two correct responses, the modulation depth for each component, m, was decreased, while following one incorrect response it was increased. The step size was 4 dB (in terms of $20 \log m$) until four reversals had occurred, after which it was decreased to 2 dB and eight more reversals were obtained. The threshold for a given run was taken as the mean value of $20 \log m$ at the last eight reversals. Each threshold reported here is based on the mean of four runs.

TABLE I. Characteristics of the three-component modulators. Values of the maximum (Max) and minimum (Min) are given relative to the rms value of the ac component of the modulator for each modulator type and each frequency spacing of the components.

Frequency spacing	Type I		Type II		Type III	
Hz	Max	Min	Max	Min	Max	Min
5	2.43	−2.43	1.81	−1.81	2.36	−2.44
10	2.37	−2.37	1.78	−1.78	2.13	−2.44
25	2.03	−2.03	1.79	−1.79	1.07	−2.44
40	2.45	−2.45	1.67	−1.67	2.44	−2.44
45	2.43	−2.43	1.81	−1.81	2.36	−2.44

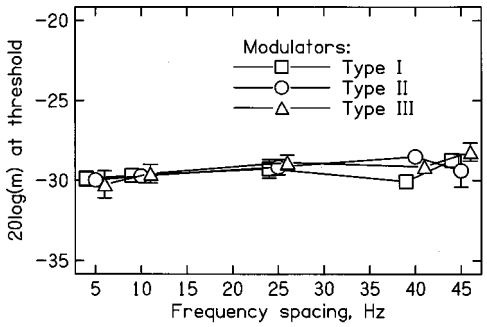


FIG. 2. Mean results across four subjects for experiment 1. Modulation detection thresholds (expressed as the value of the modulation index, *m*, for each component, converted to dB), are plotted as a function of the frequency spacing of the components. The frequency of the center component was always 50 Hz. Each symbol indicates a different modulator type, as indicated in the key. Error bars indicate ± one standard error. Error bars are omitted where they would be smaller than the symbol used to represent a particular condition.

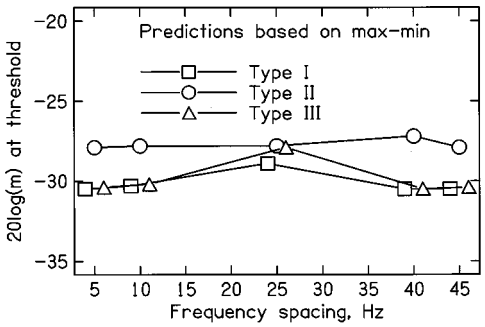


FIG. 3. Predictions of the results of experiment 1 based on the assumption that threshold corresponds to a fixed value of the envelope maximum to the envelope minimum (max–min).

C. Subjects

Four subjects were tested. One was author AS. The other three subjects were paid for their services. All subjects had absolute thresholds less than 20 dB HL at all audiometric frequencies and had no history of hearing disorders. All had previous experience in psychoacoustic tasks. They were given at least 10 h practice, after which performance appeared to be stable. The thresholds gathered during the practice sessions were discarded.

D. Results

The results were very similar across subjects, and the mean results are shown in Fig. 2. Error bars show ± one standard error. Thresholds are expressed as the modulation index, *m*, for each component of the three-component modulator, converted to dB ($20 \log m$). It is clear that the mean thresholds are very similar for the three modulator types (phase conditions). Furthermore, there is no clear effect of the frequency spacing of the modulator components. A within-subjects analysis of variance (ANOVA) was conducted with factors modulator type and frequency spacing of the modulator components. Neither the main effects nor the interaction were statistically significant.

The results are somewhat surprising. We had expected that, at least for the smaller frequency spacings, the modulation detection thresholds would be influenced by the relative phase of the components. As noted in the Introduction, previous work had suggested that thresholds might be related to the max–min statistic (Strickland and Viemeister, 1996) or to the crest factor of the envelope (Lorenzi *et al.*, 1999), and, for a fixed value of *m*, our stimuli differed in max–min value and crest factor across modulator type.

To illustrate the expected effects of manipulating relative modulator phase, we generated predictions based on the max–min value and the crest factor. In other words, we calculated what the thresholds should be if each threshold corresponded to a constant value of the max–min or the crest factor. The steps in doing this for the max–min value were as follows. First, we calculated the grand mean threshold across conditions, which was −29.33 dB, in terms of $20 \log m$. Then, for each condition (each modulator type and frequency spacing) we calculated the max–min value that would occur for that condition for a modulation depth of −29.33 dB. The max–min values were averaged across conditions, giving a value mean(max–min). Then, for each condition, we calculated what modulation depth would be required to give the value mean (max–min). This value is the predicted threshold for that condition. The predictions generated in this way are shown in Fig. 3. Generally, thresholds for the type I and type III modulators were predicted to be similar, while thresholds for the type II modulator were predicted to be 2–3 dB higher. Higher thresholds for the type II modulator were not observed in the experimental results. The ANOVA indicated that the standard error of differences between mean thresholds for the different modulator types was only 0.36 dB. Thus, differences of 2–3 dB should easily have been measured. In fact, the mean obtained thresholds for the different modulator types were almost identical, at −29.37, −29.35, and −29.83 dB for types I, II, and III, respectively.

Figure 4 shows predictions calculated in a similar way for the crest factor. For the frequency spacings of 5, 10, 40, and 45 Hz, the predictions are similar to those for the max–min value. For the frequency spacing of 25 Hz, the threshold is predicted to be markedly higher for the type III modulator than for the other modulators, as, for a fixed value of *m*, the crest factor was very small for the type III modulator (see Table I). This pattern of results was not observed in the experimental data.

In summary, the pattern of our results is not consistent

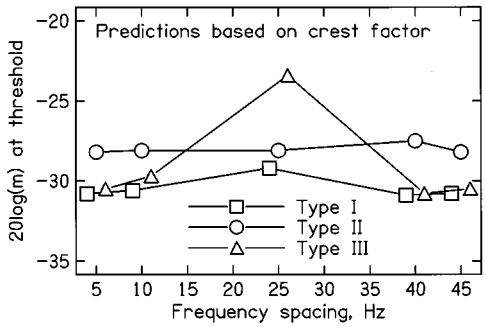


FIG. 4. Predictions of the results of experiment 1 based on the assumption that threshold corresponds to a fixed crest factor of the envelope (ratio of peak to rms value).

with the idea that thresholds correspond to a constant max–min value or a constant crest factor. Rather, thresholds were almost independent of relative modulator phase, and corresponded to a constant rms modulation depth. As there was no effect of relative modulator phase even for the small frequency spacings, the results cannot be used to assess the validity of the concept of the MFB.

III. EXPERIMENT II: PSYCHOMETRIC FUNCTIONS FOR MODULATION DETECTION

Although the results of experiment 1 showed no significant effect of relative modulator phase, the three different modulator types did lead to slightly different modulation detection thresholds for the two largest frequency spacings (40 and 45 Hz). This might indicate a small effect of modulator phase. Therefore, we decided to increase the precision of the results by measuring psychometric functions for modulation detection, for the three modulator types.

A. Method

A two-alternative, forced-choice procedure with feedback was used, as described earlier. The method for determining psychometric functions was the same as described by Moore and Sek (1992, 1994a, 1994b); the reader is referred to those papers for details. Briefly, each run of 55 trials started with five trials where the signal contained a highly detectable amount of modulation, so that subjects ''knew what to listen for.'' Scores for these five trials were discarded. In subsequent trials, five different modulation depths were used, in a repeating sequence going from larger values to smaller ones. Thus, subjects received a ''reminder'' (easy) stimulus once every five trials. Each point on each psychometric function was based on 200 judgments. The stimuli were essentially the same as for experiment 1, except that only three frequency spacings were used: 5, 25, and 45 Hz. Three of the subjects from experiment 1 were used, including author AS.

B. Results

The percent-correct scores were converted to d' values using standard tables (Hacker and Ratcliff, 1979). Scores for each condition and each subject are shown in Fig. 5. Since the detectability of AM is roughly proportional to m^2 (Moore and Sek, 1992, 1994a, 1994b), the d' values are plotted as a function of $1000m^2$ (the factor of 1000 is included simply to give convenient numbers). It is clear that the psychometric functions are similar across subjects, and are similar for the different modulator types and frequency spacings. An ANOVA was conducted on the d' values with factors modulation depth, modulator type, and frequency spacing. The main effect of modulation depth was highly significant: $F(4,8)=328.1$, $p<0.001$. The other main effects were not significant, and none of the interactions was significant. We conclude that the detectability of the AM was not significantly influenced by the relative phase of the modulator components or by the frequency spacing of the modulator components.

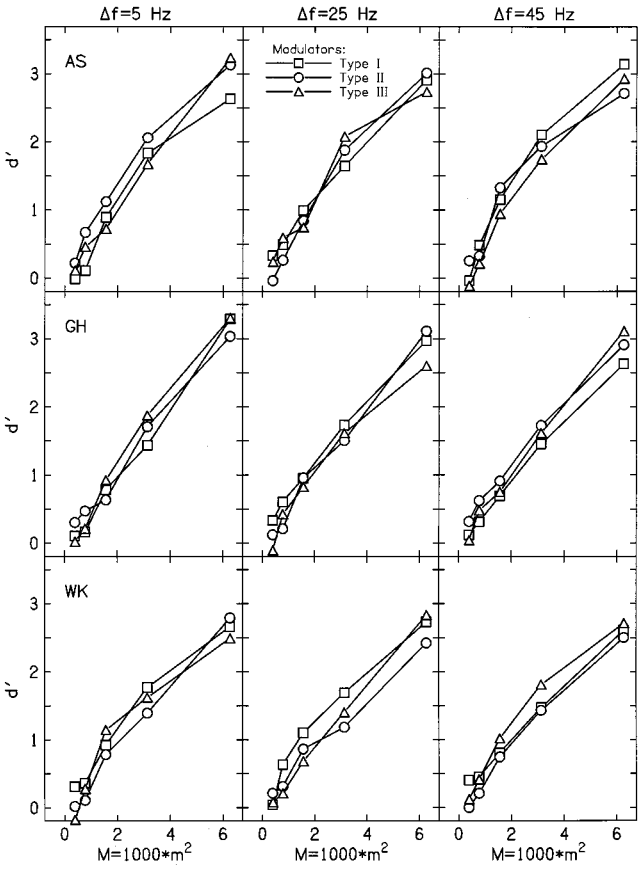


FIG. 5. Results of experiment 2. Detectability (d') is plotted as a function of $1000m^2$ with modulator type as parameter. Each row shows results for one subject, and each column shows results for one frequency spacing of the components.

If detectability depended on the max–min ratio or on the crest factor, then detectability averaged across frequency spacings should have been lower for the type II modulator than for types I or III, for the same reasons as described earlier. In fact, there was no hint of this in the data: mean d' values (averaged across both modulation depth and frequency spacing) were 1.20, 1.18, and 1.17 for modulator types I, II, and III, respectively. Also, for the type III modulator, detectability should have been lower for the 25-Hz spacing than for the other spacings, especially if the crest factor were the decision variable. Again, this was not the case. Values of d' averaged across modulation depth for the type III modulator were 1.19, 1.13, and 1.18, for the 5, 25, and 45-Hz spacings, respectively. These findings support the conclusion of experiment 1, that detectability does not depend on the max–min value or the crest factor.

IV. DISCUSSION

Our results appear to be discrepant with those of Strickland and Viemeister (1996) and Lorenzi *et al.* (1999), although it must be acknowledged that their tasks were very different from ours. Strickland and Viemeister (1996) found that their results were consistent with the idea that the detectability of signal AM in the presence of masker AM is determined by the max–min value of the envelope. Lorenzi *et al.* (1999) proposed that the detectability of changes in

relative modulator phase is determined by changes in the crest factor of the envelope. Our results indicate that the detectability of modulation is not determined by either the max–min value or the crest factor.

It is possible that different aspects of the envelope are used to perform different tasks; the aspect applicable to absolute thresholds for modulation detection may be different from the aspect applicable to masked thresholds for modulation detection or to the detection of changes in relative modulator phase. However, it is worth considering whether the results of all three tasks can be accounted for within a common framework, using a common decision statistic. In particular, we now examine the possibility that the discrepancy between the different experiments is related to the different modulation depths used. Strickland and Viemeister used masker modulation that was highly detectable ($m = 0.25$ or 0.5) and determined the signal modulation depth needed for detection of the signal modulation. Lorenzi *et al.* used two-component modulators with the modulation index of one fixed at 0.5. For a modulator frequency ratio of 2:1, changes in relative modulator phase could be heard when the modulation index of the other component was 0.25 or more, but not when it was less than 0.25. In our experiment, the modulation depth of all modulator components was low (less than 0.08).

It seems likely that the internal representations of the envelopes of stimuli with large modulation depths are ''distorted'' by various nonlinearities in the auditory system (Shofner *et al.*, 1996). Examples of such nonlinearities include the compressive nonlinearity on the basilar membrane for midrange sound levels (Sellick *et al.*, 1982; Robles *et al.*, 1986; Ruggero *et al.*, 1997), and adaptation and saturation effects in the auditory periphery (Smith, 1979; Shofner *et al.*, 1996). We previously suggested that the nonlinearity associated with basilar-membrane compression could explain results obtained using ''beating'' two-component modulators, arguing that the distortion of the modulator waveform introduced an envelope component at the beat rate (Moore *et al.*, 1999a). We proposed that this distortion component could account for the modulation masking observed for a modulation signal at the beat rate, and could account for the observed effects of the phase of that signal relative to the envelope of the beating masker modulator. Verhey *et al.* (2000) have also presented evidence for an envelope component at the beat rate of two-component modulators, although the pattern of their phase effects was different from that observed by Moore *et al.* (1999a). They argued that that a nonlinearity other than basilar-membrane compression was responsible for the distortion in the internal representation of the envelope.

The distortion of the modulator waveform produced by nonlinearities in the auditory system would be small for low modulation depths, but greater for large modulation depths. It is possible that the modulation detection threshold depends on the rms value (or the mean-squared value) of the envelope *after* it has passed through the nonlinearities. For *small* modulation depths, modulators with equal rms modulation depth before passing through the nonlinearities would have roughly equal rms modulation depths after passing through the nonlinearities, regardless of the modulator waveform. Thus, equal rms modulation depths would lead to equal detectability, consistent with our results. For larger modulation depths, modulators with equal rms values but with different max–min values or different crest factors would have different rms values after passing through the nonlinearities.

To examine this issue in a more quantitative way, we used the simple model described by Moore *et al.* (1998, 1999b). The envelopes of the stimuli were passed through a compressive nonlinearity, resembling the compressive input–output function of the basilar membrane. We do not want to argue that this is the only nonlinearity involved. We use it here merely to illustrate that the general pattern of the results can be accounted for with this form of nonlinearity. The nonlinearity had the following form:

$$L_{\text{out}} = 0.9L_{\text{in}} + 6 + 53(1 - (1/(1 + \exp(-0.05(L_{\text{in}} - 50))))), \quad (1)$$

where L_{in} is the input level in dB and L_{out} is the output level in dB; see Fig. 4 in Moore *et al.* (1999b). This function is based on an equation suggested by Oxenham (1995), and used subsequently by Moore *et al.* (1996) and by Oxenham and Moore (1997). Various statistics were computed based on the envelope before and after compression. For the stimuli used in the present experiment, the rms value of the ac component of the envelope (equal to the standard deviation, s.d., of the envelope) was reduced by the compression, but the s.d. after compression was not influenced by relative modulator phase. This is consistent with our finding that the relative modulator phase did not influence the detection thresholds.

We also applied this model to some of the stimuli of Strickland and Viemeister (1996). Some caution is needed here, as they used a noise carrier, and the model is based only on the average envelope of the stimuli, and ignores the inherent random amplitude fluctuations present for a noise carrier. The model may not be realistic in this case, as the inherent fluctuations of a noise carrier clearly have an influence on modulation detection (Dau *et al.*, 1997b, 1997a, 1999). However, Strickland and Viemeister used relatively large modulation depths for their sinusoidal modulation masker, so it seems reasonable to assume that the intrinsic fluctuations in the noise played only a minor role in this case.

We started by considering a case where the masker modulator had a modulation index of 0.5 and the masker frequency was twice the signal frequency. When the signal modulation depth was set to -20 dB (corresponding to $m = 0.1$, a ''representative'' value when the signal modulation was at threshold; the change in envelope power, the measure used by Strickland and Viemeister, is 0.005), the s.d. of the envelope after compression was affected by the relative phase of the signal and masker. For example, the s.d. was greater when the signal starting phase (relative to that of the masker) was 135° than when it was 45°. For several different relative starting phases of the signal, we calculated the modulation index of the signal needed to produce a fixed s.d. after compression. These calculated values represent predictions of the detection threshold of the signal based on the assumption that threshold corresponds to a constant s.d. after

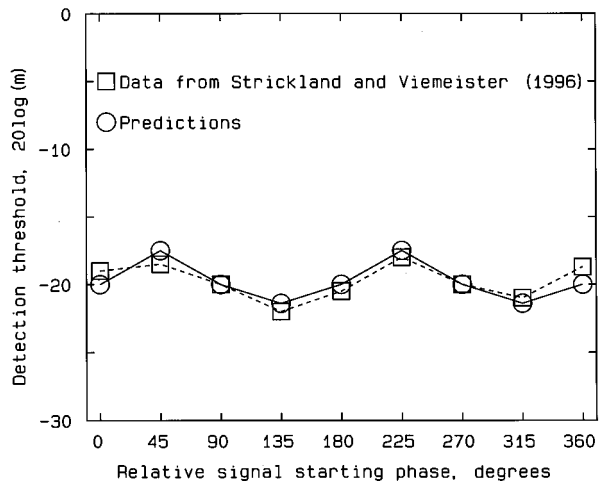


FIG. 6. Predictions of modulation detection thresholds for a signal modulator in the presence of a masker modulator with twice the frequency of the signal, based on the assumption that threshold corresponds to a constant rms value of the ac component of the envelope after compression on the basilar membrane. Predicted thresholds (circles) are plotted as a function of the starting phase of the signal relative to the masker. Mean results obtained by Strickland and Viemeister (1996) for a signal frequency of 16 Hz and a masker frequency of 32 Hz are shown by squares.

compression. The predictions are shown by the circles in Fig. 6. Thresholds are predicted to be lowest (best) for relative phases of 135° and 315° and highest for relative phases of 45° and 225°. For comparison, the squares show the mean data obtained by Strickland and Viemeister for a masker frequency of 32 Hz and a signal frequency of 16 Hz. The match between the predictions and the data is better than for any of the models evaluated by Strickland and Viemeister.

In other cases, the model did not fit the data of Strickland and Viemeister (1996). When the masker frequency is half the signal frequency, the model predicts best (lowest) thresholds for a relative phase of 90°, and worst (highest) thresholds around 270°; the predictions depend only on the masker/signal frequency ratio. The predicted phase effects are close to the observed effects for a masker frequency of 50 Hz and a signal frequency of 100 Hz, but are almost opposite to those obtained for a masker frequency of 8 Hz and a signal frequency of 16 Hz. This discrepancy suggests that the compressive nonlinearity on the basilar membrane is not the only nonlinearity involved. A time-varying nonlinearity, perhaps connected with adaptation effects, would be needed to explain the pattern of results observed by Strickland and Viemeister (1996).

As noted in the Introduction, several previous researchers have proposed that the rms value of the envelope is used as a decision statistic for modulation detection and discrimination. For example, Wakefield and Viemeister (1990) used the leaky integrator model of Viemeister (1979) in a attempt to model their data on modulation depth discrimination, using the rms value of the envelope (extracted using a broad initial filter, a half-wave rectifier, and a low-pass filter) as a decision statistic. They found that the model accounted for their results quite well for small standard modulation depths, but that predicted thresholds were lower (better) than obtained thresholds for large modulation depths. However, they did not include any nonlinearity prior to calculating the rms value. Possibly, a better fit to the data would be obtained if such a nonlinearity were included.

The simple model described above does not adequately account for the frequency selectivity that has been observed in the modulation domain. For example, Strickland and Viemeister (1996) measured thresholds for detecting sinusoidal signal modulation as a function of the sinusoidal masker modulation frequency, keeping the modulation depth of the masker fixed. For signal frequencies of 4 and 16 Hz, the thresholds were highest (worst) when the masker frequency equaled the signal frequency, and decreased progressively with increasing masker–signal frequency separation (although somewhat ''jagged'' masking functions were obtained for two subjects when the masker frequencies included values that were not harmonics or subharmonics of the signal frequency). Predicted thresholds based on the rms value after compression did not show the observed pattern. Thresholds were predicted to be higher when the masker and signal frequencies were equal than when they were unequal, but the predicted thresholds did not change systematically with increasing masker–signal frequency separation. The model and decision statistics used by Strickland and Viemeister also did not predict the data well. It may be necessary to assume an MFB, or a sensitivity to an aspect of the modulator other than the rms value after application of a nonlinearity, to account for frequency selectivity in the modulation domain.

One other simplification in our analysis should be noted. We have considered only the response of a single ''channel'' centered at the signal frequency. Probably, modulation detection and discrimination involve the integration of information from several frequency channels. For the analysis of our data, the simplification is probably irrelevant. For the small modulation depths used in our experiment, equal rms depths of the input signal would lead to equal rms depths at the output of the model for any frequency channel (although the rms value after passing through one or more nonlinearities would differ across channels). Also, as the overall level was fixed, the same frequency channels would have been involved in all conditions. For experiments using a noise carrier, such as those of Strickland and Viemeister (1996) and Lorenzi *et al.* (1999), the modulation after passing through any nonlinearity would have been similar in all channels, except to the extent that basilar-membrane compression appears to be greater for high frequencies than for low frequencies (Hicks and Bacon, 1999; Moore *et al.*, 1999c). Thus, it may be reasonable to consider the response of a single channel. However, this simplification may lead to serious errors when considering modulation detection data for sinusoidal signals whose overall level is varied (Riesz, 1928; Wojtczak and Viemeister, 1999). For such data, the number of channels conveying useful information increases with increasing level, and this probably plays a role in the decrease of modulation detection thresholds with increasing level (Florentine and Buus, 1981; Moore and Sek, 1994b).

In summary, our results are consistent with the idea that modulation detection threshold depends on the rms value of the envelope *after* it has been subjected to one or more non-

linearities in the auditory system. Some aspects of the results of Strickland and Viemeister (1996) are also consistent with this idea, although their data suggest that at least one component of the nonlinearity is time varying.

V. CONCLUSIONS

(1) The detectability of a three-component modulator applied to a 4000-Hz sinusoidal carrier was not influenced by the relative phase of the components or the frequency spacing of the components.

(2) The results are not consistent with the idea that modulation detection threshold corresponds to a fixed value of the max–min value (the ratio of the maximum envelope amplitude to the minimum envelope amplitude) or to a fixed value of the crest factor.

(3) The results are consistent with the idea that modulation detection threshold corresponds to a fixed rms value of the ac component of the envelope after application of one or more nonlinearities. Basilar-membrane compression is one such nonlinearity, but others, such as adaptation, may play a role.

(4) It may be necessary to assume an MFB, or a sensitivity to an aspect of the modulator other than the rms value after application of nonlinearities, to account for frequency selectivity in the modulation domain.

ACKNOWLEDGMENTS

This work was supported by the Wellcome Trust and by the MRC. We thank Tom Baer for considerable assistance with modeling and for helpful comments on an earlier version of this paper. We also thank Torsten Dau and an anonymous reviewer for very helpful comments and Neal Viemeister for helpful discussions.

Bacon, S. P., and Grantham, D. W. (**1989**). ''Modulation masking: Effects of modulation frequency, depth and phase,'' J. Acoust. Soc. Am. **85**, 2575–2580.

Dau, T. (**1996**). ''Modeling auditory processing of amplitude modulation,'' Ph.D. thesis, University of Oldenburg.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997a**). ''Modeling auditory processing of amplitude modulation. I. Detection and masking with narrowband carriers,'' J. Acoust. Soc. Am. **102**, 2892–2905.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997b**). ''Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration,'' J. Acoust. Soc. Am. **102**, 2906–2919.

Dau, T., Verhey, J. L., and Kohlrausch, A. (**1999**). ''Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers,'' J. Acoust. Soc. Am. **106**, 2752–2760.

Festen, J. M., and Plomp, R. (**1981**). ''Relations between auditory functions in normal hearing,'' J. Acoust. Soc. Am. **70**, 356–369.

Florentine, M., and Buus, S. (**1981**). ''An excitation-pattern model for intensity discrimination,'' J. Acoust. Soc. Am. **70**, 1646–1654.

Forrest, T. G., and Green, D. M. (**1987**). ''Detection of partially filled gaps in noise and the temporal modulation transfer function,'' J. Acoust. Soc. Am. **82**, 1933–1943.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hear. Res. **47**, 103–138.

Hacker, M. J., and Ratcliff, R. (**1979**). ''A revised table of d' for M-alternative forced choice,'' Percept. Psychophys. **26**, 168–170.

Hartmann, W. M., and Pumplin, J. (**1988**). ''Noise power fluctuations and the masking of sine signals,'' J. Acoust. Soc. Am. **83**, 2277–2289.

Hicks, M. L., and Bacon, S. P. (**1999**). ''Psychophysical measures of auditory nonlinearities as a function of frequency in individuals with normal hearing,'' J. Acoust. Soc. Am. **105**, 326–338.

Houtgast, T. (**1989**). ''Frequency selectivity in amplitude-modulation detection,'' J. Acoust. Soc. Am. **85**, 1676–1680.

Lorenzi, C., Berthommier, F., and Demany, L. (**1999**). ''Discrimination of amplitude-modulation phase spectrum,'' J. Acoust. Soc. Am. **105**, 2987–2990.

Moore, B. C. J., Glasberg, B. R., Plack, C. J., and Biswas, A. K. (**1988**). ''The shape of the ear's temporal window,'' J. Acoust. Soc. Am. **83**, 1102–1116.

Moore, B. C. J., Launer, S., Vickers, D., and Baer, T. (**1998**). ''Loudness of modulated sounds as a function of modulation rate, modulation depth, modulation waveform and overall level,'' in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London).

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1996**). ''Detection of decrements and increments in sinusoids at high overall levels,'' J. Acoust. Soc. Am. **99**, 3669–3677.

Moore, B. C. J., and Sek, A. (**1992**). ''Detection of combined frequency and amplitude modulation,'' J. Acoust. Soc. Am. **92**, 3119–3131.

Moore, B. C. J., and Sek, A. (**1994a**). ''Discrimination of modulation type (AM or FM) with and without background noise,'' J. Acoust. Soc. Am. **96**, 726–732.

Moore, B. C. J., and Sek, A. (**1994b**). ''Effects of carrier frequency and background noise on the detection of mixed modulation,'' J. Acoust. Soc. Am. **96**, 741–751.

Moore, B. C. J., Sek, A., and Glasberg, B. R. (**1999a**). ''Modulation masking produced by beating modulators,'' J. Acoust. Soc. Am. **106**, 908–918.

Moore, B. C. J., Vickers, D. A., Baer, T., and Launer, S. (**1999b**). ''Factors affecting the loudness of modulated sounds,'' J. Acoust. Soc. Am. **105**, 2757–2772.

Moore, B. C. J., Vickers, D. A., Plack, C. J., and Oxenham, A. J. (**1999c**). ''Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism,'' J. Acoust. Soc. Am. **106**, 2761–2778.

Oxenham, A. J. (**1995**). ''Psychophysical consequences of peripheral auditory nonlinearity,'' Ph.D. thesis, Cambridge.

Oxenham, A. J., and Moore, B. C. J. (**1994**). ''Modeling the additivity of nonsimultaneous masking,'' Hear. Res. **80**, 105–118.

Oxenham, A. J., and Moore, B. C. J. (**1997**). ''Modeling the effects of peripheral nonlinearity in listeners with normal and impaired hearing,'' in *Modeling Sensorineural Hearing Loss*, edited by W. Jesteadt (Erlbaum, Hillsdale, NJ).

Plack, C. J., and Moore, B. C. J. (**1990**). ''Temporal window shape as a function of frequency and level,'' J. Acoust. Soc. Am. **87**, 2178–2187.

Riesz, R. R. (**1928**). ''Differential intensity sensitivity of the ear for pure tones,'' Phys. Rev. **31**, 867–875.

Robles, L., Ruggero, M. A., and Rich, N. C. (**1986**). ''Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input-output functions, tuning curves, and response phases,'' J. Acoust. Soc. Am. **80**, 1364–1374.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (**1997**). ''Basilar-membrane responses to tones at the base of the chinchilla cochlea,'' J. Acoust. Soc. Am. **101**, 2151–2163.

Sellick, P. M., Patuzzi, R., and Johnstone, B. M. (**1982**). ''Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique,'' J. Acoust. Soc. Am. **72**, 131–141.

Shofner, S., Sheft, S., and Guzman, S. J. (**1996**). ''Responses of ventral cochlear nucleus units in the chinchilla to amplitude modulation by low-frequency, two-tone complexes,'' J. Acoust. Soc. Am. **99**, 3592–3605.

Smith, R. L. (**1979**). ''Adaptation, saturation and physiological masking in single auditory-nerve fibers,'' J. Acoust. Soc. Am. **65**, 166–178.

Strickland, E. A., and Viemeister, N. F. (**1996**). ''Cues for discrimination of envelopes,'' J. Acoust. Soc. Am. **99**, 3638–3646.

Verhey, J. L., Ewert, S., and Dau, T. (**2000**). ''Modulationsverarbeitung im Gehör: Autokorrelation versus Filterbank,'' DAGA 2000 (in press).

Viemeister, N. F. (**1979**). ''Temporal modulation transfer functions based on modulation thresholds,'' J. Acoust. Soc. Am. **66**, 1364–1380.

Wakefield, G. H., and Viemeister, N. F. (**1990**). ''Discrimination of modulation depth of SAM noise,'' J. Acoust. Soc. Am. **88**, 1367–1373.

Wojtczak, M., and Viemeister, N. F. (**1999**). ''Intensity discrimination and detection of amplitude modulation,'' J. Acoust. Soc. Am. **106**, 1917–1924.

Failure to unlearn the precedence effect

R. Y. Litovsky,[a] M. L. Hawley, and B. J. Fligor
Hearing Research Center and Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215

P. M. Zurek
Sensimetrics Corporation, Somerville, Massachusetts 02144 and Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 19 July 1999; revised 22 May 2000; accepted 1 August 2000)

Studies of the precedence effect using two binaural clicks have shown that listeners' ability to discriminate changes in the interaural time difference (ITD) of the lagging click is much poorer than that for the leading click [e.g., Zurek, J. Acoust. Soc. Am. **67**, 952–964 (1980)]. This difference is thought to reflect an auditory process that suppresses directional information from the lagging sound and attributes greater perceptual weight to directional information contained in the leading one. A report by Saberi and Perrott [J. Acoust. Soc. Am. **87**, 1732–1737 (1990)] suggested that listeners can ''unlearn'' this suppression of the lag's directional information after training with an adaptive psychophysical procedure involving 100 reversals and extremely small step sizes. Here, an attempt was made to find a similar effect using psychophysical procedures that are more common to precedence studies. Eight subjects were rigorously trained on the precedence task using either a blocked procedure or an adaptive procedure to vary ITD. Listeners showed no sign of unlearning. After 9–31 h of participating in the task, all subjects maintained high lag just-noticeable differences (jnd's) and low single source jnd's. This failure to train away the precedence effect (as manifested in discrimination suppression) suggests that directional information contained in the lagging source is not easily accessed. Several possible explanations for the discrepancies between the present study and Saberi and Perrott's finding are discussed.
[S0001-4966(00)01311-4]

PACS numbers: 43.66.Qp, 43.66.Pn [DWG]

I. INTRODUCTION

People spend an abundance of time in enclosed spaces, such as homes, classrooms, and work environments. A common feature of these environments is their reverberant nature due to the hard surfaces comprising the walls, floors, and ceiling, as well as furniture and objects contained in the room. In a typical classroom, the initial reflections that reach the ears from nearby surfaces may be attenuated by as little as 5–10 dB relative to the direct sound. Hence, the listener receives a complex mixture of acoustic signals, each carrying its own localization cues. In order to avoid localization errors, the auditory system must resolve those cues belonging to the source and weight them more heavily in the localization process.

Efforts to understand how the auditory system processes sounds in complex environments have utilized simple stimulus paradigms in which a direct sound (lead) and a single simulated reflection (lag) are presented in anechoic environments (or over headphones) with short delays between their onsets. ''Summing localization'' refers to a delay smaller than 1 ms, when the lead and lag sources are perceptually fused and when both the lead and lag contribute to the perceived direction of the fused image (e.g., Zurek, 1980; for review see Blauert, 1997, pp. 204–206; Litovsky *et al.*, 1999). As the delay is increased to 1 ms and beyond, the lead dominates localization. Fifty years ago, Wallach, Newman, and Rosenzweig (1949) coined the term ''precedence effect'' to denote the localization dominance of the leading sound. The dominance of the lead in localization has also been shown using an objective task whereby listeners discriminate changes in the ITD or location of the lag. Compared with discrimination of the lead or a single source, lag discrimination is significantly more difficult for clicks at delays of 1–3 ms (Zurek, 1980; Gaskell, 1983; Freyman *et al.*, 1991; Shinn-Cunningham *et al.*, 1993; Litovsky and Macmillan, 1994; Stellmack *et al.*, 1998; Tollin and Henning, 1998; Litovsky *et al.*, 1999). This finding of lag discrimination suppression provides an objective measure of the extent to which the auditory system suppresses directional cues contained in the lag and attributes greater perceptual weight to directional information contained in the leading source (Zurek, 1980; Litovsky *et al.*, 1999).

A number of years ago, Saberi and Perrott (1990) published a report summarizing group data from three listeners that indicated that lag discrimination suppression can be ''trained away'' with extensive practice. Using brief click stimuli, they initially measured just-noticeable-difference thresholds (jnd's) in the lag condition that revealed relatively poor performance at lead–lag delays ranging from 0.75–3 ms, consistent with previous reports (e.g., Zurek, 1980; Gaskell, 1983). Listeners were subsequently ''trained'' on the precedence task at delays ranging from 1–5 ms with emphasis at 2.35 ms for 8–20 h. A curious, and perhaps

[a]Electronic mail: litovsky@bu.edu

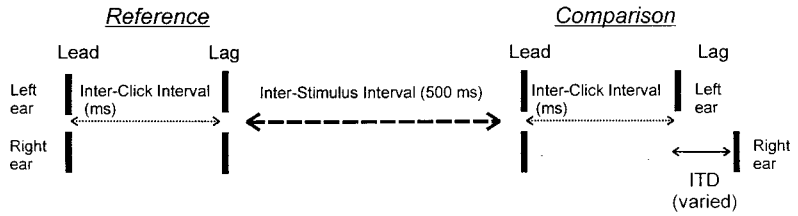


FIG. 1. Stimulus configuration for a single trial is shown. Each click was a 40-μs square-wave pulse. Each trial consisted of two intervals: A reference stimulus and a comparison stimulus, which were separated by 500 ms. Each stimulus consisted of two pairs of clicks, the lead and lag. The lead pair was always diotic (zero interaural-time difference producing an image centered in the middle of the head). The lag pair was diotic in the reference stimulus as well, but dichotic in the comparison stimulus. The interclick interval (ICI) was always identical for the reference and comparison, and varied from 0.38 to 10 ms.

critical, feature of the psychophysical procedure was the implementation of unusually long adaptive tracks with very small step sizes. Following training, ITD discrimination on the lag sound improved substantially, although single-source jnd's were not reported; hence, it is difficult to surmise whether the training effect was specific to precedence or a more general improvement in ITD discrimination. Nonetheless, Saberi and Perrott (1990) concluded that listeners learned to use the directional cues available in the lag, and that the precedence effect results from a cognitive process that is subject to modification from short-term experience. The notion that precedence mechanisms are not ''hard-wired'' and can be unlearned negates most existing models of precedence which invoke active inhibition in order for the dominance of the lead to be activated (e.g., Lindermann, 1986; Zurek, 1987; Cai *et al.*, 1998).

The approach taken by Saberi and Perrott (1990) to estimate threshold used unusually long adaptive tracks that continued for 100 reversals with ITD step sizes of 2.5 μs for the last 60 reversals. This method has not been previously used in studies of the precedence effect; hence, the ''unlearning'' effect may not apply to most current research. Virtually all studies on precedence have employed alternative strategies that either average over several shorter adaptive runs (e.g., Litovsky and Macmillan, 1994; Yang and Grantham, 1997; Tollin and Henning, 1998) or that estimate psychometric functions with numerous measures at fixed ITDs (e.g., Zurek, 1980; Gaskell, 1983; Yost and Soderquist, 1984; Perrott *et al.*, 1989; Freyman *et al.*, 1991; Shinn-Cunningham *et al.*, 1993; Houtgast and Aoki, 1994; Stellmack *et al.*, 1999). The present study explored the extent to which the unlearning effect extends to these more common psychophysical methods. If unlearning precedence can be produced independent of a particular psychophysical method, existing models of the precedence effect would have to be modified in fundamental ways. Alternatively, if the unlearning effect is dependent on a specific method of tracking subject's performance, then much would remain to be understood about the particulars of this finding.

II. TRAINING WITH FIXED ITDs

A. Methods

1. Subjects

Five listeners (two male and three female) participated in this study. All listeners had normal hearing as verified by audiograms measured at standard frequencies ranging from 250 to 8000 Hz. One listener (S5) was a student with prior experience in auditory psychophysics, but not on tasks requiring ITD discrimination; the remainder of the listeners were paid subjects with little or no prior experience. Listeners were familiarized with the task for approximately 1 h prior to testing.

2. Stimuli

Square-wave pulses of 40-μs duration were presented as lead–lag binaural click pairs. Figure 1 shows the temporal sequence of stimuli on a single trial. Each trial consisted of two intervals. Interval A, containing the reference stimulus, and interval B, containing the comparison stimulus, were separated by 500 ms. Each stimulus consisted of two pairs of clicks, the lead and lag. The lead pair was diotic (ITD$=0$) in both the reference and comparison. The lag pair was diotic in the reference, and contained an ITD favoring either the right or left ear in the comparison stimulus. The interclick interval (ICI), referring to the delay between onsets of the lead and lag, was always identical for the reference and comparison. The task was a two-alternative, forced-choice task, whereby the listener was asked to report whether the sound image in the comparison stimulus was perceived to the right or left of center. Feedback was provided on every trial. Tucker-Davis Technologies' AP2 array processor and PD1 Power DAC were used to generate the stimuli. The sampling rate was 200 kHz, thereby allowing interaural differences as small as 5 μs. Stimuli were amplified by a TDT HB6 headphone amplifier and presented at a level of 68 dB SPL[1] via Sennheiser HD520II headphones.

In this first experiment we aimed to provide listeners with directional cues associated with the lag that were as consistent as possible. We chose to maintain consistency by holding the lag ITD constant within a block of trials and only varying the direction of change (right vs left). Our hope was that as listeners became more familiar with the task they would be better able to utilize information regarding changes in the lag ITD.

B. Results

1. Pretraining assessment

Pretraining thresholds were obtained using a 2-down/1-up adaptive procedure (Levitt, 1971) to vary the ITD. Starting ITDs were 500 μs and each run continued until 14

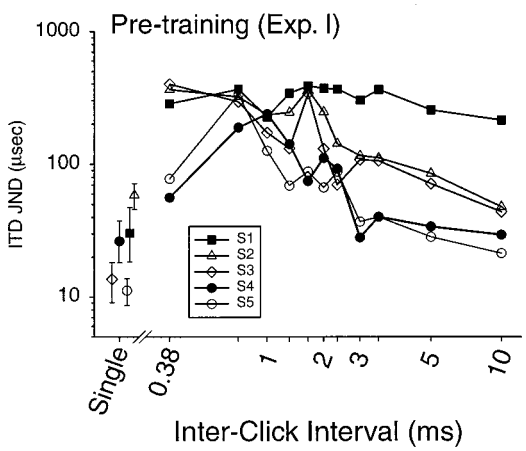


FIG. 2. Fixed training group pretraining ITD jnd's for lag discrimination. Individual lines denote data from individual listeners, with performance plotted as a function of interclick interval (lead–lag delay). Single points on the left represent individuals' performance on the no-lead (single-source) condition.

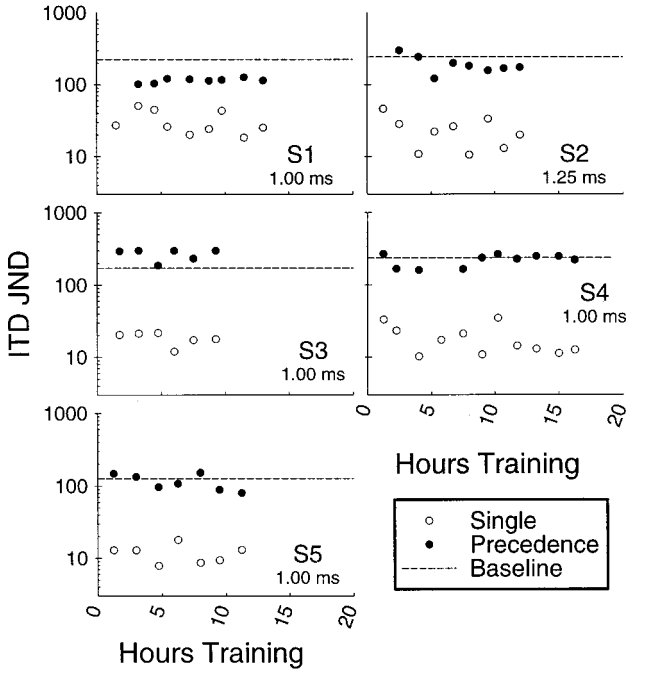


FIG. 3. Daily thresholds for the fixed training group are shown. Each panel contains data from one listener, with ITD jnd's for each day plotted as a function of number of training hours completed by the end of that day. Filled circles refer to performance on lag discrimination, and open circles refer to the no-lead (single source) condition. The value shown under the subject number refers to the interclick interval (ICI) at which the listener was trained. Horizontal dashed lines refer to thresholds obtained in the pretraining period for the ICI at which the listener was trained.

reversals were reached, with threshold for each run calculated from the geometric mean of the last ten reversals. ITDs were increased or decreased by a multiplicative factor of 2 for the initial four reversals, and by a factor of 1.4 for the last ten reversals. Measurements were made at 11 ICIs (0.38, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.5, 3, 5, and 10 ms), and each listener was tested on five series of ICIs, with order of presentation randomized differently for each series. These pretraining estimates were gathered over approximately four 2-h sessions.

Pretraining data for the fixed training group are shown in Fig. 2. Solid curves show performance on the precedence lag-discrimination condition, and single points at the left show performance on the no-lead (single-source) condition for individual listeners. For three subjects (S1, S4, S5) lag discrimination is poorest at delays of 0.75 to 3 ms and performance generally improves as the delays either increase (5–10 ms) or decrease (0.38 ms). Two subjects (S2, S3) do not show the expected improvement at the shortest delay, possibly due to the difficulty of the task, which may have rendered neither lead nor lag dominant at that time interval. Finally, to ensure that subjects were not simply ''unlearning'' the precedence effect during the pretraining assessment, a two-way analysis of variance (ANOVA) (threshold repetition×ICI) was conducted for each subject. There were no significant main effects or interactions ($p>0.05$).

2. Training with a blocked ITD procedure

Training was conducted during 6 to 11 daily sessions, which lasted 9 to 17 h. Based on these initial measurements for each listener (Fig. 2), an ICI near the poorest ITD jnd thresholds was selected as the training ICI (1.0 ms for S1, S2, S3, and S5; 1.25 ms for S2). Although 1.0 ms is considered to be at the edge of the ''summing localization'' window, the fact that performance was worse at that ICI for most listeners convinced us that for these listeners and conditions the precedence effect was well in place. On each day, one psychometric function was obtained at the training ICI by measuring discrimination performance for blocks of 50 trials with the ITD fixed within a block. ITDs were chosen semi-randomly between 50 and 800 μs, and measurements were repeated three times at each ITD; hence, data are based on 150 trials per ITD. Daily thresholds seeking the ITD at which performance was 70.7% correct were obtained from psychometric functions by linear interpolation between the nearest points above and below 70.7. In addition, every day, prior to testing on the precedence conditions, three single-source baseline thresholds were obtained using the adaptive 2-down/1-up method (Levitt, 1971).

Figure 3 shows individual ITD jnds' plotted as a function of training hours. Lag-discrimination thresholds are shown in filled circles and single-source thresholds are shown in open circles. The dashed line in each plot marks the listener's baseline performance at the training ICI (see average ITD jnd at that ICI in Fig. 2). Results suggest that lag-discrimination thresholds remained consistently above single-source thresholds across hours of training. There was no evidence of improvement in either condition as listeners gained more experience on the task.

III. TRAINING WITH ADAPTIVE CHANGES IN ITD

The methods used in the fixed training study differed from those of Saberi and Perrott (1990) in several ways. For instance, they had used an adaptive method to vary ITDs, and their training emphasized an ICI of 2.35 for all listeners. We wondered whether listeners' ability to unlearn precedence was dependent on either the adaptive changes in ITD

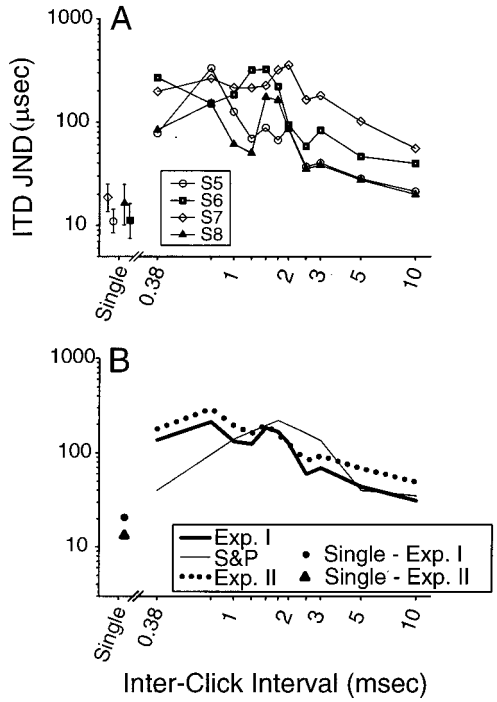


FIG. 4. (A) Adaptive training group pretraining ITD jnd's for lag discrimination. Individual lines denote data from individual listeners, with performance plotted as a function of interclick interval (lead–lag delay). Single points on the left represent individuals' performance on the no-lead (single-source) condition. (B) Group means from (A) and from Fig. 2 are shown together with the group mean from Saberi and Perrott (1990, S&P).

as a function of performance, or on the fact that we chose a training ICI near the peak of each listener's function. We decided to conduct a second experiment in which the lag ITDs varied adaptively within a run, and to fix the training ITD at 2.35 ms.

A. Method

1. Subjects

Four subjects (two male and two female) participated. One listener (S5) had participated in experiment 1 and the other listeners had no prior experience in auditory psychophysics. All listeners had normal hearing as verified by audiograms measured at standard frequencies ranging from 250 to 8000 Hz. Listeners were familiarized with the task for approximately 1 h prior to testing.

B. Results

1. Pretraining assessment

Pretraining thresholds were obtained using a method identical to that for the first experiment; hence, results are expected to be similar to those in Fig. 2. Data for the adaptive method are shown in Fig. 4. The top panel (A) shows four listeners' performance on lag-discrimination as a function of ICI, and single points at the left show individual's performance on the no-lead (single-source) condition. Lag discrimination is generally poorest at the shortest delays and performance generally improves as the delays increase. The expected finding that performance also improves at delays less than 1 ms (see Blauert, 1997) was only clearly observed for S5.

The bottom panel in Fig. 4 compares group mean data from Saberi and Perrott's (1990) pretraining condition (thin line) to group mean data from our fixed (dotted line) and adaptive (thick line) training groups. It appears that in both of our groups poor performance in the precedence condition was observed at delays as short as 0.38 ms. In contrast, Saberi and Perrott's (1990) subjects had a clearer region of poor performance with improvement not only at the longer delays but at delays less than 1 ms as well. This difference is difficult to interpret, but might be due to the difficulty of the task, which extended to short delays for our subjects. Finally, as in the first experiment, to ensure that subjects were not simply unlearning the precedence effect during the pretraining assessment, a two-way ANOVA (threshold repetition$\times$ICI) was conducted for each subject. There were no significant main effects or interactions ($p > 0.05$).

2. Training with an adaptive method

Training was conducted during 7 to 15 daily sessions, which corresponded to 13 to 23 h. As in the fixed training group, number of hours refers to actual time spent listening to stimuli, not including intermittent breaks taken within a session. Training was conducted at an ICI of 2.35 ms for all listeners. Note that for three listeners (S5, S6, and S8) this ICI was greater than their pretraining peak jnd (see Fig. 4), and for one listener (S7) 2.35 was near the peak jnd. One listener (S8) showed excellent performance at 2.35 ms, before training, and after 11 h of training was subsequently tested at an ICI of 1.75 ms, near the peak jnd. Every day, prior to testing on the precedence conditions, one single-source baseline threshold was obtained using the adaptive 2-down/1-up method. Subsequently, testing was conducted for the precedence lag-discrimination condition. ITD jnd thresholds were measured adaptively using a PEST algorithm (Taylor and Creelman, 1967; Hawley, 1994). For both single and precedence runs the starting ITDs were 500 μs. On average, 35 thresholds were measured per day (one single source and the remainder precedence). Stimuli were identical to those used in the first experiment.

Figure 5 shows daily performance for individual listeners. Each plot shows the mean and standard deviation bars for the lag-discrimination thresholds (filled circles), and single-source thresholds (open circles), as a function of training time. The horizontal dashed line in each plot marks the listener's baseline performance at the training ICI [see the average ITD jnd at that ICI in Fig. 4(A)]. Three listeners (S5, S6, and S7) exhibited a consistent elevated threshold on the precedence condition compared with the single condition. For these listeners, precedence thresholds remained stable at the dashed line, suggesting that with increased training they were not learning to extract information regarding the ITD of the lag at smaller ITDs. At the same ICI of 2.35 ms, one listener (S8) exhibited better performance on the precedence condition compared with the pretraining baseline performance; however, lag-discrimination thresholds remained

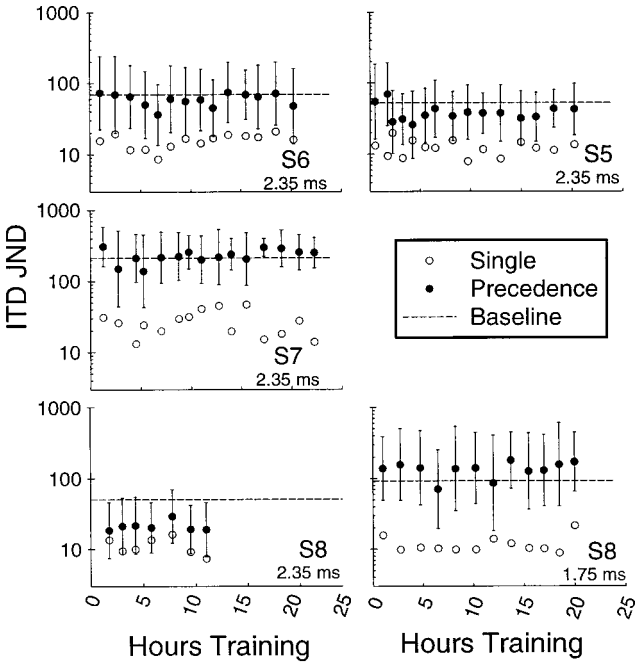


FIG. 5. Daily thresholds for the adaptive training group are shown. Each panel contains data from one listener, with daily average and standard deviation for ITD jnd's. Filled circles refer to performance on lag discrimination, and open circles refer to the no-lead (single-source) condition. The value shown under the subject number refers to the interclick interval (ICI) at which the listener was trained; this value was initially 2.35 for all listeners. S8 was also tested at an ICI of 1.75 (bottom right panel). Horizontal dashed lines refer to thresholds obtained in the pretraining period for the ICI at which the listener was trained.

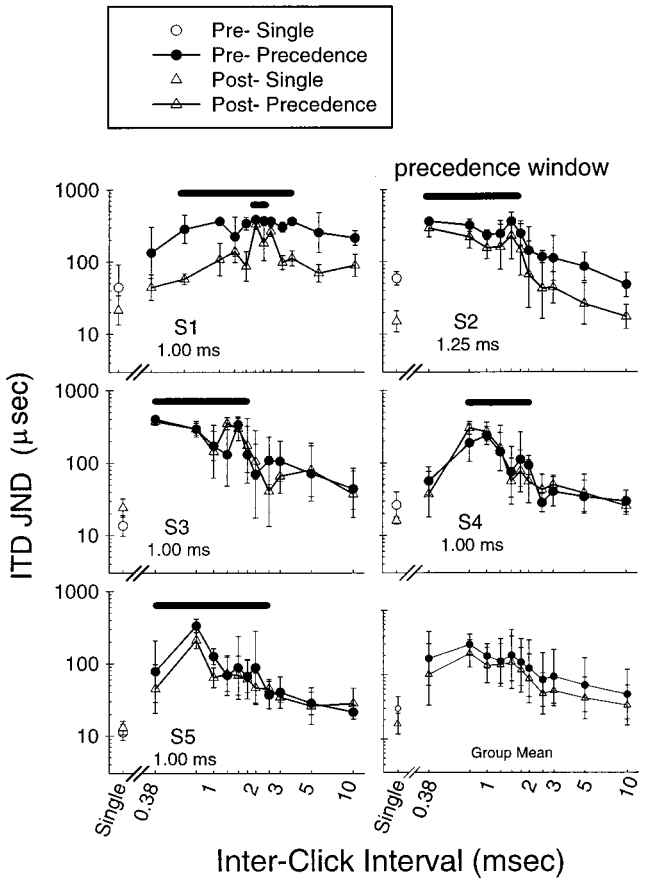


FIG. 6. Comparing pre- and post-training performance for the fixed training. Each panel contains data from one listener, with ITD jnd's plotted as a function of the interclick interval (ICI). Pre-training data are shown by circles and post-training data are shown by triangles. Single points on the left represent data from the single-source (no lead) condition. Horizontal bars across the top of each plot mark the ''precedence window,'' defined as the range of delays at which performance was significantly higher than on the single-source condition.

consistently higher than the single-source thresholds after 11 h of training. This listener was subsequently trained at an ICI of 1.75 ms, near the pretraining peak jnd. Data for the 1.75-ms ICI are also shown in Fig. 5 (bottom right). Again, after 20 additional training hours, lag-discrimination thresholds remained consistently near baseline performance, and greater than single-source thresholds.

IV. COMPARISON OF PRE- AND POST-TRAINING FOR EXPERIMENTS 1 AND 2

Following the training period, each listener was once again tested on the baseline ICI series, with five adaptive measures at each ICI. Figure 6 shows pre- and post-training thresholds as a function of ICI for each listener in the fixed training group (experiment 1) with group means at the bottom right. In each plot, pretraining performance is shown in circles, and post-training in triangles. Data from the single-source condition are also shown in each plot as a single data point on the left. Figure 7 shows pre- and post-training thresholds for the adaptive training group (experiment 2). Finally, group means are plotted in Fig. 8 for the fixed training group (A), adaptive training group (B), and for Saberi and Perrott's listeners (panel C).

To analyze the single-source results, for each listener a one-way ANOVA was conducted using the five repeated estimates obtained before and after training. There was a significant main effect for two listeners, although pointing towards opposite conclusions. For S2 pretraining thresholds were higher than post-training thresholds ($p<0.001$), while for S3 pretraining thresholds were lower than post-training thresholds ($p<0.05$).

To analyze the lag-discrimination thresholds, a two-way ANOVA was conducted for each listener, comparing pre- and post-training performance across ICIs, including the single-source condition as one of the ICI values. All listeners showed a main effect of ICI ($p<0.0001$), suggesting a dependence of precedence performance on the delay between lead and lag. Statistical tests also showed one listener to have a main effect of pre-post condition (S2, $p<0.0001$). Finally, one listener (S1) had a significant interaction between ICI and pre-post condition (S1, $p<0.009$), revealing an effect of training only at ICIs that were outside the region of 0.75 to 3 ms ($p<0.05$). For this listener, small amounts of improvement were seen either at ICIs of 0.0, 0.2, 0.38, or at ICIs beyond the classic precedence window (5 and 10 ms). However, this listener showed no improvement within the traditional precedence window of 1 to 3 ms for clicks (e.g., Zurek, 1980; Gaskell, 1983; Saberi and Perrott, 1990; Tollin and Henning, 1998).

The next analysis was aimed at identifying a ''prece-

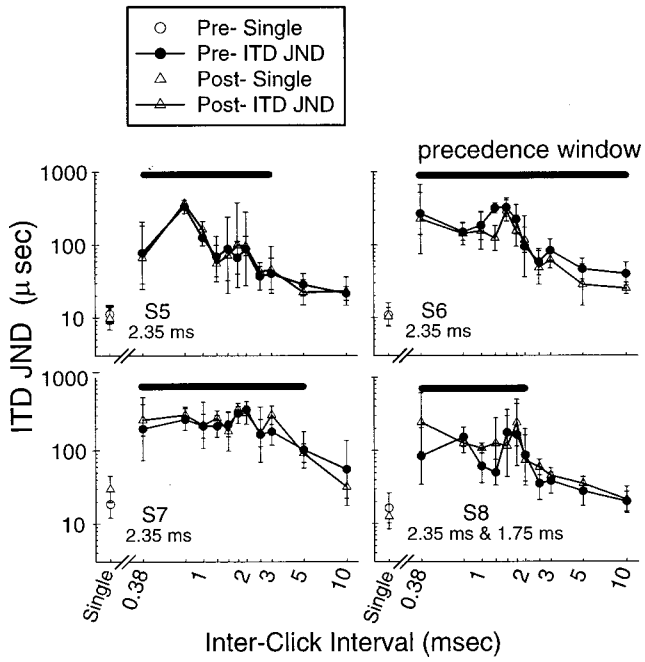


FIG. 7. Same as Fig. 6 for the adaptive training group.

dence window'' for each listener, defined as a temporal region for which lag discrimination is significantly worse than single-source discrimination. Subsequent to the general ANOVAs, *post hoc* Scheffe's grouping tests were conducted for each listener to reveal the ICIs that fell into the precedence window. At the top of each plot in Fig. 6, thick horizontal bars indicate the precedence window for each listener. For listeners with no effect of pre-post training, thresholds were combined within a given ICI. Listener S1, for whom there was an interaction effect between ICI and training, the precedence window was smaller in the post-training sessions (1.5 and 1.75 ms) compared with a more extensive precedence window in the pretraining sessions (0.75 to 10 ms). For this listener, pretraining precedence thresholds were significantly higher than single-source thresholds at most ICIs, whereas post-training thresholds were only higher than single source at two ICIs. For S2 there was no difference in the precedence window for pre- and post-training. Three listeners (S1, S4, and S5) showed a ''peaked'' function, with a significant decrease in thresholds at both short and long ICIs. For the other listeners (S2, S3, S6, S7, and S8) the precedence window extended to the very short ICIs, within the range of summing localization (e.g., Blauert, 1997).

In summary, comparison of pre- and post-training shows that, with a few minor exceptions, training did not result in improvement on the precedence task, at the training ICI or at any other ICIs. Overall, after training on the precedence task for 13 to 31.5 h performance remained constant. In addition, the precedence window, defining a range of ICIs at which precedence was most robust, varied among listeners. The precedence windows measured here are in agreement with previous studies showing strong precedence at delays of 1–3 ms (e.g., Zurek, 1980; Gaskell, 1983; Shinn-Cunningham *et al.*, 1993; Tollin and Henning, 1998; Stellmack *et al.*, 1998).

V. DISCUSSION

This study was aimed at finding an effect of unlearning precedence similar to one reported by Saberi and Perrott (1990). We began with a commonly used method of fixing the interaural parameters of the lagging source. After failing to reduce our listeners' lag-ITD thresholds following 9 to 17 h of training, we switched to another common paradigm whereby the ITD of the lag is varied adaptively. After 13 to 23 h of training we still saw no improvement in performance, and thus concluded that the precedence effect is not easily unlearned and must therefore have some hardwired mechanisms that are rapid, automatic, and not cognitively mediated.

The discrepancy in findings must seriously be considered, however, since Saberi and Perrott's unlearning effect was robust enough to be carried over to tone pips of various frequencies and high-pass-filtered clicks. We considered the possibility that Saberi and Perrott did not actually demonstrate a release from precedence, but rather a more general learning effect that simply reflected listeners' improvement on the interaural discrimination task. If their listeners had shown improvement on all ITD thresholds, including those with a single click, then an argument could be made that the training effect was not specific to ITD discrimination on the lagging click. However, because they did not report single-click thresholds, it is not possible to assess this hypothesis directly. It is reasonable to assume, though, that there would have been little if any training effect for single-click ITD thresholds in their study because their listeners were moderately-to-very experienced with lateralization tasks. This supposition is supported by the results of their first study in which lead-click ITD thresholds were about 20 μs. Therefore, we regard it as highly unlikely that Saberi and

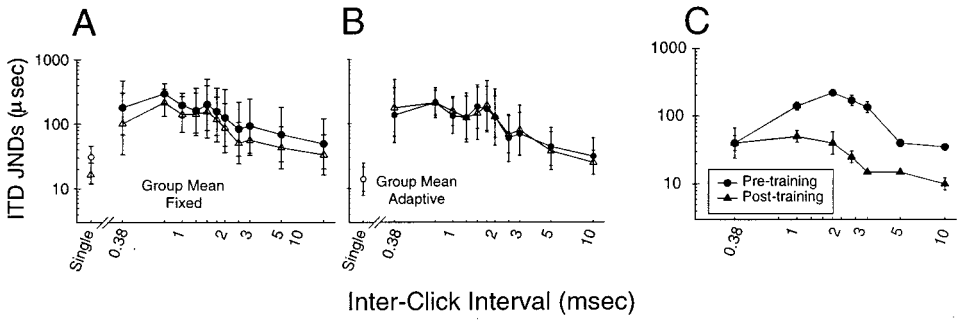


FIG. 8. Group means for the fixed training method (A) and adaptive training method (B); ITD jnd thresholds are plotted as a function of ICI for pre-training (○) and post-training (△) sessions. (C) Data from Saberi and Perrott (1990) are replotted also showing ITD jnd's as a function of ICI.

Perrott's training result was a general learning effect.

Why is it that we failed to find the effect of unlearning the precedence effect? Is it possible that their highly trained subjects may have had been better able to listen for the appropriate cues for discriminating ITD of the lagging click? Evidence against this hypothesis comes from the fact that all three of Saberi and Perrott's listeners, even the one with the least experience, showed the training effect while our listeners, even the one with the most experience, did not. Further, almost all of the published data on lag-ITD discrimination suppression (e.g., Zurek, 1980; Shinn-Cunningham *et al.*, 1993; Saberi, 1996; Tollin and Henning, 1998; Litovsky and Shinn-Cunningham, 2000) have employed highly trained listeners but have not reported unlearning of the kind shown by Saberi and Perrott.

Our failure to find the effect may in fact be due to methodological and procedural details. First, in Saberi and Perrott's (1990) study, listeners were trained each day using a single long adaptive track with linear step sizes beginning with 10 μs, and ending with 2.5 μs. In our study, subjects were trained either with fixed ITDs or with repeated adaptive tracks using PEST rules. In either case, subjects in their study spent more time listening to trials near thresholds, which might have facilitated improved discrimination and lowering of thresholds. Second, while they used a 300-ms interstimulus interval, we used a 500-ms interval. The shorter interval used by Saberi and Perrott may have contributed to an ''apparent motion'' percept that could conceivably increase sensitivity to directional changes in the comparison stimulus compared with the reference stimulus. Third, in the training phase we focused on the delay at which each listener's threshold had been highest in the pretraining phase, with the aim of presenting listeners with consistent relatively unchanging interaural cues, which we thought would maximize any effect of training. In contrast, Saberi and Perrott varied the delay within the 1–5-ms window of interest, which might have actually been a better paradigm for inducing a general learning effect. Finally, the level in the present study was higher, leaving the possibility that we had weaker precedence (Shinn-Cunningham *et al.*, 1993; Litovsky and Yin, 1998). That might explain our failure to find a training effect if it were not for the fact that thresholds were lower (i.e., weaker precedence) in the Saberi and Perrott study. However, nothing is known about the effect of sound level on training, and hence the impact of a level difference on the discrepancy in results cannot be established.

Although these methodological differences exist between the present study and that of Saberi and Perrott, it is difficult at this point to conceive of learning mechanisms that would be sensitive to such seemingly inconsequential details.

VI. SUMMARY

For a number of years the phenomenon of lag-discrimination suppression has been a widely documented, objectively measurable, manifestation of the precedence effect (Zurek, 1980; Gaskell, 1983; Freyman *et al.*, 1991; Shinn-Cunningham *et al.*, 1993; Litovsky and Macmillan, 1994; Stellmack *et al.*, 1998; Tollin and Henning, 1998; Litovsky and Shinn-Cunningham, 2000; see Litovsky *et al.*, 1999 for a recent review). Some authors have argued that an active inhibitory mechanism is necessary to account for this and related precedence phenomena, and that a likely candidate for initial stages of this process lies in the brainstem (Lindemann, 1986; Zurek, 1987; Yin, 1994; Cai *et al.*, 1998; Litovsky and Yin, 1998). However, with one exception (Fitzpatrick *et al.*, 1995), most of the physiological data supporting these claims have been collected in anesthetized animals. At the same time, it is clear that higher-order mechanisms must be involved in many of the precedence phenomena that have been shown to be susceptible to changes in listeners' ongoing auditory experience and environmental awareness (e.g., Hafter *et al.*, 1988; Clifton *et al.*, 1994; Clifton and Freynamn, 1997; Blauert, 1997). Whether the loss of ITD discrimination ability shortly after a preceding sound is malleable in the same way is an important piece of knowledge for a basic understanding of auditory processing. However, before we can develop sound theoretical models, we need to establish the basic experimental findings and understand the source of differences in findings.

ACKNOWLEDGMENTS

This work was supported by NIH Grants No. R29-DC03083 to R. Y. Litovsky and ONR MURI Award No. Z883402. Monica Hawley was supported by NIH Grant No. R01-DC00100. A version of these data appeared in an abstract at the Association for Research in Otolaryngology meeting, Feb. 1999.

[1]For these brief stimuli, defining the actual level in dB SPL depends on the time window chosen for the measurement. Over a 1000-ms time window, which is equivalent to the ''slow'' setting on a sound-level meter, the level was 68 dB SPL. Using a shorter time window of 35 ms, which is equivalent to the ''impulse'' setting on a sound-level meter, the measured level was 79 dB SPL.

Blauert, J. (**1997**). *Spatial Hearing: The Psychophysics of Human Sound Localization*, revised ed. (The MIT Press, Cambridge, MA).

Cai, H., Carney, L. H., and Colburn, H. S. (**1998**). ''A model for binaural response properties of inferior colliculus neurons. I. A model with ITD-sensitive excitatory and inhibitory inputs,'' J. Acoust. Soc. Am. **103**, 475–493.

Clifton, R. K., and Freyman, R. L. (**1997**). ''The precedence effect: Beyond echo suppression,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Erlbaum, Mahwah, NJ), pp. 233–255.

Clifton, R. K., Freyman, R. L., Litovsky, R. Y., and McCall, D. (**1994**). ''Listeners' expectations about echoes can raise or lower echo threshold,'' J. Acoust. Soc. Am. **95**, 1525–1533.

Fitzpatrick, D. C., Kuwada, S., Batra, R., and Trahiotis, C. (**1995**). ''Neural responses to simple, simulated echoes in the auditory brainstem of the unanesthetized rabbit,'' J. Neurophysiol. **74**, 2469–2486.

Freyman, R. L., Clifton, R. K., and Litovsky, R. Y. (**1991**). ''Dynamic processes in the precedence effect,'' J. Acoust. Soc. Am. **90**, 874–884.

Gaskell, H. (**1983**). ''The precedence effect,'' Hear. Res. **12**, 277–303.

Hafter, E. R., Buell, T. N., and Richards, V. (**1988**). ''Onset-coding in lateralization: Its form, site, and function,'' in *Auditory Function: Neurobiological Bases of Hearing*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 647–676.

Hawley, M. L. (**1994**). ''Comparison of adaptive procedures for obtaining psychophysical thresholds using computer simulation,'' M.S. thesis, Boston University, Biomedical Engineering.

Houtgast, T., and Aoki, S. (**1994**). ''Stimulus-onset dominance in the perception of binaural information,'' Hear. Res. **72**, 29–36.

Levitt, H. (**1971**). ''Transformed up–down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Lindemann, W. (**1986**). ''Extension of a binaural cross-correlation model by contralateral inhibition. II. The law of the first wavefront,'' J. Acoust. Soc. Am. **80**, 1623–1630.
Litovsky, R. Y., and Macmillan, N. A. (**1994**). ''Sound localization precision under conditions of the precedence effect: Effects of azimuth and standard stimuli,'' J. Acoust. Soc. Am. **96**, Pt. 1, 752–758.
Litovsky, R. Y., and Yin, T. C. T. (**1998**). ''Physiological studies of the precedence effect in the inferior colliculus of the cat. I. Correlates of psychophysics,'' J. Neurophysiol. **80**, 1285–1301.
Litovsky, R. Y., Colburn, H. S., Yost, W. A., and Guzman, S. J. (**1999**). ''The precedence effect,'' J. Acoust. Soc. Am. **106**, 1633–1654.
Litovsky, R. Y., and Shinn-Cunningham, B. G. (**2000**). ''Investigation of the relationship between three common measures of precedence: fusion, localization dominance and discrimination suppression,'' J. Acoust. Soc. Am. (to be published).
Perrott, D. R., Marlborough, K., Merrill, P., and Strybel, T. Z. (**1989**). ''Minimum audible angle thresholds obtained under conditions in which the precedence effect is assumed to operate,'' J. Acoust. Soc. Am. **85**, 282–288.
Saberi, K., and Perrott, D. R. (**1990**). ''Lateralization thresholds obtained under conditions in which the precedence effect is assumed to operate,'' J. Acoust. Soc. Am. **87**, 1732–1737.
Saberi, K. (**1996**). ''Observer weighting of interaural delays in filtered impulses,'' Percept. Psychophys. **58**(7), 1037–1046.
Shinn-Cunningham, B. G., Zurek, P. M., and Durlach, N. I. (**1993**). ''Adjustment and discrimination measurements of the precedence effect,'' J. Acoust. Soc. Am. **93**, 2923–2932.
Stellmack, M. A., Dye, R. H., and Guzman, S. J. (**1998**). ''Observer weighting of binaural information in source and echo clicks,'' J. Acoust. Soc. Am. **105**, 377–387.
Taylor, M. M., and Creelman, C. D. (**1967**). ''PEST: Efficient estimates on probability functions,'' J. Acoust. Soc. Am. **41**, 782–787.
Tollin, D. J., and Henning, G. B. (**1998**). ''Some aspects of the lateralization of echoed sound in man. I. Classical interaural delay-based precedence,'' J. Acoust. Soc. Am. **104**, 3030–3038.
Wallach, H., Newman, E. B., and Rosenzweig, M. R. (**1949**). ''The precedence effect in sound localization,'' Am. J. Psychol. **LXII**(3), 315–336.
Yang, X., and Grantham, D. W. (**1997**). ''Echo suppression and discrimination suppression aspects of the precedence effect,'' Percept. Psychophys. **59**, 1108–1117.
Yin, T. C. T. (**1994**). ''Physiological correlates of the precedence effect and summing localization in the inferior colliculus of the cat,'' J. Neurosci. **14**, 5170–5186.
Yost, W. A., and Soderquist, D. R. (**1984**). ''The precedence effect: Revisited,'' J. Acoust. Soc. Am. **76**, 1377–1383.
Zurek, P. M. (**1980**). ''The precedence effect and its possible role in the avoidance of interaural ambiguities,'' J. Acoust. Soc. Am. **67**, 952–964.
Zurek, P. M. (**1987**). ''The precedence effect,'' in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer, New York), pp. 85–105.

Investigating perceptual features of electrode stimulation via a multidimensional scaling paradigm

Leslie M. Collins[a] and Chandra S. Throckmorton[b]
Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina 27708-0291

(Received 11 November 1998; revised 1 July 1999; accepted 1 August 2000)

To achieve the most effective speech processing for individuals with cochlear implants, it is important to understand the perceptual features associated with the stimulation parameters. In general, when electrodes are stimulated in order from apex to base, the pitch of the perceived sound changes in an orderly fashion from low to high. Some deviations from this assumed order have been documented. Also, pitch is the dominant perceptual attribute of a sound when the stimuli associated with different electrodes have been accurately loudness balanced. In this study, the results of a multidimensional scaling (MDS) paradigm were compared to the results of a pitch-ranking procedure for six subjects implanted with multichannel cochlear prostheses. Results indicate that there may be multiple percepts that change with electrode location. Not surprisingly, the dominant percept is strongly correlated with pitch. The results also indicate that the structure of the second percept is consistent across subjects, although not interpretable using the data measured in this study. Furthermore, results indicate that MDS data can be used to pinpoint indiscriminable electrodes more accurately than pitch data. The results of this study may have importance for the design of the next generation of speech processors. © *2000 Acoustical Society of America.* [S0001-4966(00)02511-X]

PACS numbers: 43.66.Ts, 43.66.Dc, 43.71.Ky, 43.64.Me [JWH]

I. INTRODUCTION

To achieve the most effective speech processing for individuals with cochlear implants, it is important to understand the mechanisms by which to manipulate the perceptual features associated with the parameters of stimulation as completely as possible. Pulse trains presented at a constant level on a single electrode have been reported to sound like a ''buzz,'' ''hum,'' or ''drumming sound'' on some electrodes or to produce a distinct pitch on others (Pauka, 1989). Pauka (1989) hypothesized that electrodes that produce a wider current spread would result in the ''buzzier'' percepts and those that produce narrower current spreads would produce the ''pitch-like'' percepts.

It is generally accepted that a percept that is related to the pitch dimension of acoustical stimulation is elicited by stimulation of a single electrode, and in accordance with the place–pitch theory, when pairs of electrodes are stimulated in order from the apex to the base, the pitch of the perceived sound changes in a fairly orderly fashion from low to high (Simmons, 1966; Eddington *et al.*, 1978a, 1978b, 1980; Townshend *et al.*, 1987; Tong *et al.*, 1980, 1982, 1985; Shannon, 1983a; Dowell *et al.*, 1982; Pauka, 1989; Dorman *et al.*, 1990; Soli, 1990; Nelson *et al.*, 1995; Busby *et al.*, 1994; Collins *et al.*, 1997). Furthermore, when the stimuli have been accurately loudness balanced, pitch is the dominant perceptual attribute associated with stimulation of an electrode, an assumption consistent with the tonotopic order of the cochlea. In addition, it is often assumed in a clinical setting that each sound is distinct from all others, i.e., that electrodes are completely discriminable. In addition to the location in the cochlea, the rate of electrical stimulation affects pitch perception in an analogous fashion to the effect of rate in acoustic stimulation. As the stimulation rate increases on a particular electrode, subjects report an increase in the perceived pitch of the stimuli, although this increase in pitch begins to saturate at rates of approximately 500 pps (Eddington *et al.*, 1978a, 1978b, 1980; Tong *et al.*, 1980, 1983b, 1985; Shannon, 1983a; Pfingst, 1988).

The rate–pitch, place–pitch, and discriminability assumptions form the basis of modern speech-processing techniques for multichannel cochlear implants. However, several studies have indicated that there may be additional perceptual attributes that change as a function of the location at which the electrical stimulus is applied (Eddington *et al.*, 1978a; Shannon, 1983a, 1983b; Soli, 1990; McKay *et al.*, 1996; Collins *et al.*, 1997), and that some electrode pairs do not form independent channels (Shannon, 1983b; Tong and Clark, 1985; Townshend *et al.*, 1987; Pauka, 1989; Busby *et al.*, 1994; Collins *et al.*, 1997; Zwolan *et al.*, 1997; Chatterjee and Shannon, 1998; Throckmorton and Collins, 1999). To date, most studies have either investigated the structure of the pitch percept in isolation, or have investigated the structure of the perceptual space without also obtaining estimates of specific individual percepts. Thus, no studies have yet linked psychophysically derived pitch measures with estimates of the structure of the perceptual space.

Several studies have used pitch scaling, labeling, or ranking techniques to measure the pitch associated with stimulation of each electrode pair (Tong *et al.*, 1982; Shannon, 1983a; Tong *et al.*, 1983b; Tong and Clark, 1985;

[a]Electronic mail: lcollins@ee.duke.edu
[b]Electronic mail: cst@ee.duke.edu

Townshend *et al.*, 1987; Pauka, 1989; Busby *et al.*, 1994; Nelson *et al.*, 1995). These studies have reported that the pitch percepts associated with electrical stimulation vary in a fairly orderly fashion that is consistent with a tonotopic organization for most subjects. Similar results were obtained in the three studies that examined the structure of the pitch percept in subjects implanted with the Nucleus device in detail (Busby *et al.*, 1994; Nelson *et al.*, 1995; Collins *et al.*, 1997). Nelson *et al.* (1995) concluded that as electrical stimulation is varied across place of stimulation, a unidimensional percept is elicited. This conclusion contradicts *indirect* data from several other studies where it has been suggested that a multidimensional percept is associated with place of stimulation (Eddington *et al.*, 1978; Shannon, 1983b; Soli, 1990; McKay *et al.*, 1996). It also contradicts anecdotal remarks from subjects who have indicated that parameters such as roughness/smoothness or ''fullness'' of the sounds change along with pitch as a function of electrode place (Shannon, 1983a; Collins *et al.*, 1997). If pitch and timbre precepts are not independent, such anecdotal remarks would not be unexpected.

In the Collins *et al.* (1997) study, pitch-ranking and pitch-scaling tasks were performed by the same set of subjects, and the results from the two experiments were compared. Estimates of the structure of the pitch percept across the electrode array as well as estimates of the discriminability of the electrode pairs were generated from both the pitch-ranking and the pitch-scaling data. The estimates of the pitch structure obtained from the two experimental procedures were similar, though not identical. In addition, the estimates of electrode discriminability obtained from the pitch-based tasks were not identical to the more direct electrode discrimination measures obtained on the same subjects that were reported in Zwolan *et al.* (1997). The analysis methods that were based on the standard d' approach, which requires among other things an assumption of unidimensionality (Green, 1970), tended to provide less-consistent results across tasks. In several cases, electrodes that were found to have identical pitches in the ranking task were also found to be discriminable. Thus, a multidimensional model of perception was proposed to explain the data observed in these subjects; however, this model was not analyzed experimentally. These results might also indicate that pitch scaling/ranking and electrode discrimination simply differ in their abilities to estimate the perceptual distance between electrodes due to the task that was used. However, a multipercept model is also supported by anecdotal evidence (Shannon, 1983a). In this study the author notes that subjects reported that each electrode tested had a different timbre in addition to pitch and that all subjects had difficulty judging the pitch due to the changing timbre.

McKay *et al.* (1996) and Tong *et al.* (1983a) reported that the percepts elicited by two-electrode stimulation appear to be multidimensional in nature. The goal of the study by McKay *et al.* was to evaluate the perceptual dimensions evoked by nonsimultaneous dual-electrode stimulation, but the authors did include some single-electrode stimuli in their stimulus set for comparison. A multidimensional scaling (MDS) paradigm was utilized to estimate the perceptual structure associated with the set of dual- and single-electrode stimuli. Their results indicated that the dimensions of their dual-electrode stimuli could be related to the positions of the two-component electrode pairs in the array. The scope of their study did not include performing an MDS analysis of all possible single electrode stimuli or the subset of the single-electrode stimuli tested; thus, the structure of the perceptual space for the standard set of stimuli could not be evaluated. Furthermore, although there was no evidence for a multidimensional structure for single-electrode stimuli, their studies did demonstrate that their data, where two physical dimensions were used to create the stimuli, can be effectively analyzed using MDS procedures. Although the authors hypothesized that one of the perceptual dimensions associated with their stimuli was pitch, separate measures of pitch for each of the electrode pairs were not obtained, so this hypothesis could not be directly assessed.

To date, no studies have linked measures of pitch with estimates of the structure of the multidimensional perceptual space. However, several studies have suggested that multiple percepts associated with place of stimulation may exist under electrical stimulation. It is possible that dimensions such as roughness or timbre may also be present as a result of variability in current paths (i.e., overlap in the currents as the cochlea and electrode array spiral towards the base), variability in nerve survival, and distance of the electrode array from the surviving neural elements. For example, if two different electrode pairs generated current paths that were centered on the same set of neurons but one of the electrode pairs generated currents with slightly larger spatial extent in all directions, the resulting two percepts could have the same pitch but differ in perceived ''bandwidth.'' It is also possible that the rate–pitch and place–pitch percepts could interact across the array (Tong *et al.*, 1983a), and that this interaction could generate an additional perceptual dimension. The goal of this study was to begin to resolve the ambiguities between pitch-based psychophysical procedures and discrimination-based psychophysical procedures by performing a multidimensional scaling procedure and comparing the resulting data to electrode pair discrimination and pitch data. Thus, we investigate the hypothesis that percepts other than pitch associated with site of stimulation, or electrode place, may be elicited.

II. METHODS

A. Subjects

Six postlingually deafened adults implanted with a Nucleus 22-channel device participated in the experiments comprising this study. Four of the subjects were male, two were female, and the six subjects ranged in age from 48–72 years of age. All subjects had at least 1 year of experience with their device. Demographic information for these subjects is contained in Table I. Although speech recognition scores were not used as an inclusion criterion for this study, the scores from four standard speech recognition tests are listed for the subjects in Table II.

The speech processors of two of the subjects were programmed by their clinician in bipolar+2 (BP+2) mode (S2 and S4); the remaining four were programed in BP+1 mode.

TABLE I. Demographic information for the implanted subjects.

Subject	Sex	Age (years)	Age at onset of profound deafness (years)	Age at implantation (years)	Mode of stimulation
S1	M	71	25	62	BP+1
S2	M	65	20	57	BP+2
S3	M	72	16	62	BP+1
S4	M	60	56	57	BP+2
S5	F	48	41	45	BP+1
S6	F	74	57	66	BP+1

The clinically programed stimulation mode was used when presenting stimuli to subjects. In the remainder of this paper, when a specific electrode pair is referenced, it is the more basal of the bipolar pair (e.g., for the BP+1 mode of stimulation, electrode pair 1 refers to the bipolar pair [1,3]). The 22 electrodes in the array are ordered sequentially from base (1) to apex (22). Subjects had between 15 and 20 active electrodes in their array. All possible electrodes were not used in one subjects' (S1) clinically programed MAP as a result of unpleasant sensation associated with stimulation of the first four electrodes. Data for this subject are reported with electrode 5 referred to as electrode 1, etc., since electrode 5 is the most basal electrode tested. For all other subjects, electrode numbers reported are coincident with the clinical electrode numbers. Only electrodes active in the subjects' clinically programed speech processor were used in this study.

The use of human subjects as described in this manuscript was approved by the Duke University Institutional Review Board and the National Institute on Deafness and other Communication Disorders.

B. Equipment and stimuli

Stimuli were delivered to subjects through an implanted receiver/stimulator driven by a Cochlear Corporation Spectra 22 processor connected to a Cochlear Corporation dual processor interface (DPI). Stimuli were presented in the psychophysical procedures via custom-designed software that interfaced to the device through the DPI and Cochlear Corporation IF4 board. The DPI and IF4 board are the standard clinical interface hardware supplied by Cochlear Corporation. This hardware allows the clinician to program the device and present stimuli to the subject. The psychophysical procedures were run on an IBM personal computer that also controlled the DPI. A color monitor was used to cue the subjects and display feedback information when applicable.

TABLE II. Speech recognition scores of all six subjects using five different measures. Scores are reported as percent correct.

Subject	NU6 (W)	NU6 (P)	CID	Consonants	Vowels
S1	12%	27%	25%	28%	62%
S2	8%	26%	9%	27%	44%
S3	0%	0%	0%	10%	24%
S4	34%	55%	81%	53%	59%
S5	20%	39%	56%	33%	67%
S6	18%	49%	79%	41%	84%

Subjects indicated their response on each trial by depressing the appropriate key on the computer keyboard. In all of the experimental tasks, the stimuli that were presented consisted of biphasic rectangular pulses with constant pulse duration of 200 μs per phase. Stimuli were presented at a stimulation rate of 500 pps in either BP+1 or BP+2 mode, depending on the mode in which the subjects' devices were programed (see Table I). The loudness of the stimuli was adjusted by changing the amplitude of the stimuli. The stimulus duration was 300 ms and the interstimulus interval was 500 ms.

C. Experiments

All subjects participated in five psychophysical experiments: a task to measure threshold and maximum comfortable loudness, a loudness balancing task, an electrode pair discrimination task, a pitch-ranking task, and a multidimensional scaling task.

1. Threshold and maximum comfortable level task

Threshold and loudest comfortable listening level were measured using the method of adjustment for each active electrode pair. Stimulation began at the lowest possible level, and subjects used the DPI control knob to adjust the stimulation amplitude. Subjects first adjusted the amplitude to threshold, then to loudest comfortable listening level, and the levels for each were recorded.

2. Loudness balancing task

All active electrode pairs were loudness balanced to a fixed reference electrode pair located in the center of the array. In the first step of the loudness-balance procedure, the subject adjusted the loudness of the reference electrode to a ''comfortable'' level using the DPI control knob. Next, for each active electrode pair, the subject adjusted the level of the test stimulus to match that of the reference stimulus, again using the DPI control knob. Stimulation alternated between the test and reference stimulus during the adjustment. No feedback was provided. The loudness adjustment was performed six times for each pair of electrodes. A fixed AB task was used where A was always the reference electrode pair and B the test electrode pair. The initial level of the test stimulus was roved across trials; three times the initial level was near the loudest comfortable listening level and three times the initial level was near threshold. The average value of the six adjusted levels was taken to be the loudness-balance level. Following completion of the task, the electrode pairs were stimulated sequentially to verify the accuracy of the loudness balance. Roving loudness was not considered because of the potential impact of loudness on pitch (Shannon, 1983a).

3. Electrode pair discrimination task

The electrode pair discrimination task was performed using a two-interval forced-choice (2IFC) fixed-level procedure implemented as a same/different task with 60 trials per block. It was modeled on one of the two electrode discrimination tasks implemented by Zwolan *et al.* (1997). The stimulus amplitudes for each electrode pair were set based on

TABLE III. Electrodes determined to be indiscriminable based on the electrode discrimination data, electrodes determined to be indiscriminable based on the pitch-ranking data, and a consistency metric calculated from the pitch-ranking data for each subject. Square brackets indicate that the preceding electrode cannot be discriminated from all electrodes enclosed in the square brackets.

Subject	Electrode discrimination task	Pitch-ranking task	Consistency of pitch data
S1	1-2, 2-3, 4-5, 5-6, 6-7	3-5, 4-6, 7-8, 10-12, 11-[13-16], 12-[13-15], 13-[14-15], 14-15	0.719
S2	2-3, 3-4, 4-5, 9-10, 10-11, 17-18	3-[4-5], 4-5, 8-12, 10-[11,19], 12-13, 14-19, 15-[16-18], 16-[17-18], 17-18	0.733
S3	2-3, 4-5, 6-7, 7-8, 10-11, 11-12, 12-13, 13-14, 14-15, 16-17, 18-19	1-2, 4-5, 6-7, 8-13, 9-[12,14], 10-[11,16], 11-16, 12-14, 18-19	0.640
S4	2-[3,4], 3-[4-5]	1-2, 1-6, 2-6, 3-5, 3-7, 3-8, 4-5, 5-8, 6-7, 7-8, 9-10, 11-13, 12-[13-15], 14-15, 17-18, 18-19	0.703
S5	3-7, 10-12, 12-13, 14-15	1-[3-5], 3-[4-5], 6-7, 9-[10-12], 10-[11-13], 11-[12-13, 15-16], 12-[13, 15-16], 13-[14-16], 14-[15-16], 15-16, 17-18	0.737
S6	10-11	2-3, 3-4, 4-[5,7], 5-[6-7], 6-7, 8-9, 10-[11-14], 11-[12-13], 12-[13-14], 13-14, 16-17, 18-19, 19-20	0.567

the results of the loudness-balance task. All possible pairs of electrodes were tested. Discriminable electrode pairs ($d' > 1$) were tested once, indiscriminable electrode pairs were tested twice. Full feedback was provided.

4. Pitch-ranking task

A paired comparison procedure was performed to obtain the pitch-ranking data. This task was modeled on that presented by Nelson *et al.* (1995) and Collins *et al.* (1997). All possible pairs of active electrode pairs were presented to the subjects in random order, and the subjects were instructed to select the stimulus with the higher pitch. The order of the electrode pairs was randomized across intervals. The stimulus amplitudes for each electrode pair were set based on the results of the loudness-balance task. Subjects completed between 10 and 30 blocks, depending on the time available for testing. Data are presented as the average of the final ten blocks. Data are scored as the percent of the time a given electrode pair was reported to be higher in pitch than all of the other electrode pairs (percent wins). The discriminability index, d', was calculated from stimulus response matrices using the row-sum formulation originally described by David (1988) and summarized for this task in Collins *et al.* (1997).

Briefly, stimulus response matrices were calculated based on subjects' responses to all of the combinations of stimuli presented for all blocks of trials. The (i,j) entry thus corresponds to the number of times electrode i was reported by the subject to be higher in pitch than electrode j. A score a_i is determined for the ith electrode as the sum of the scores in the ith row of the stimulus response matrix, and corresponds to the number of times the ith electrode was perceived to be higher in pitch than other electrodes in the array. The row sums were then normalized to a ''percent wins'' statistic based on the total number of trials for each comparison. To generate a measure of electrode discriminability, an approach known as the least-significant difference method was employed to determine the minimum difference between two row sum scores that is required for the difference to be statistically significant (David, 1988).

5. Multidimensional scaling task

The multidimensional scaling task that was implemented for this study follows McKay *et al.* (1996) and Schiffman *et al.* (1981). Prior to data collection, all stimuli were presented sequentially to the subjects. All possible pairs of stimuli were presented in random order, and the order of the electrodes was randomized across intervals. To avoid the effects that stimulation order has on the qualities of the percepts elicited by electrical stimulation, the pairs of electrodes were presented to the subject with either the apical electrode stimulated first or the basal electrode stimulated first in a particular block of trials. Subjects completed equal numbers of ''apical first'' and ''basal first'' blocks. The subject indicated the size of the difference between the two stimuli by positioning a cursor on a bar on the computer screen. The left-hand side of the bar was labeled ''exactly the same'' and the right-hand side of the bar was labeled ''most different.'' The position was converted to a number between 0 and 1, and is thus proportional to the ''dissimilarity'' or distance between the two stimuli. At least one training block was performed prior to data collection, and four or six blocks were obtained from each subject. Dissimilarity matrices were computed by averaging the dissimilarity judgments obtained for each pair of stimuli across all trials, thus, the (i,j) and (j,i) entries correspond to the average dissimilarity between electrode pair i and electrode pair j. The resulting dissimilarity matrices were analyzed with SYSTAT (SYSTAT 5.05 for Windows, SPSS for Windows, 1994, Chicago: SPSS, Inc.) The results are analyzed and presented as the stress versus the number of dimensions in the solution, R^2 versus number of dimensions in the solution, and plots showing locations of the electrode pairs in each dimension. Stress is a measure of fit defined as the sum of the squares of the differences between distances and disparities divided by the sum of the squares of the distances and R^2 is the proportion of the optimally scaled data which is accounted for by the model.

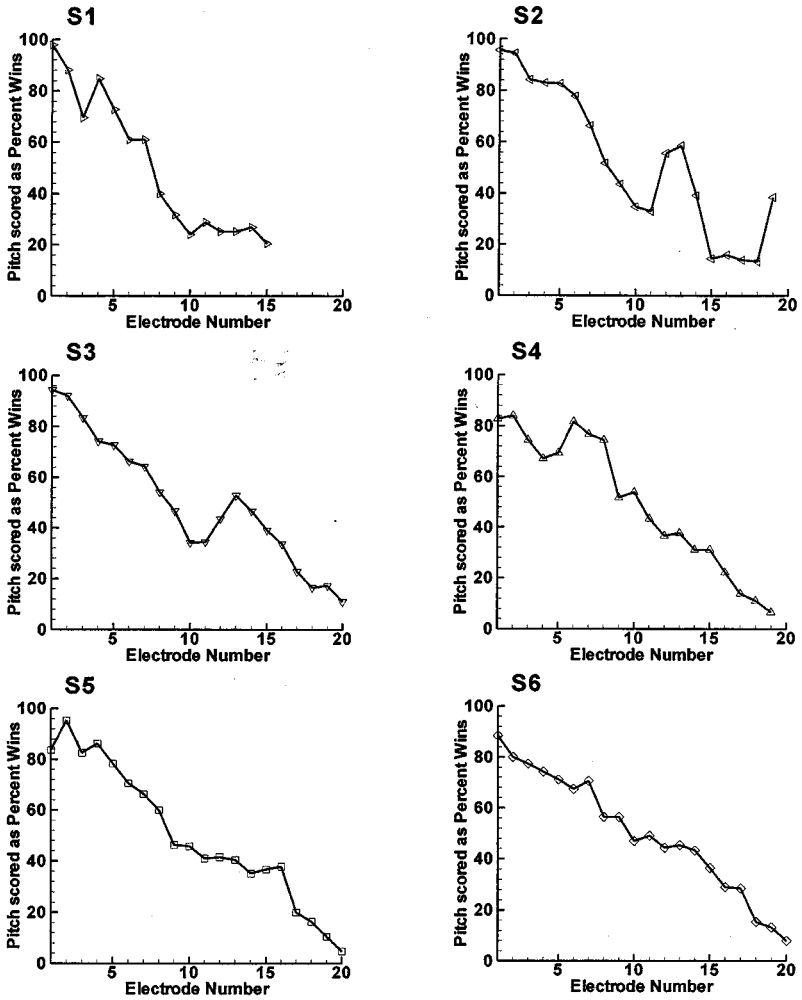


FIG. 1. Percent wins data for the pitch-ranking task for each of the six subjects who participated in this study. Percent wins are plotted on the ordinate, electrode number along the abscissa.

III. RESULTS

A. Electrode pair discrimination task

The second column of Table III lists the indiscriminable electrode pairs for each of the six subjects on the electrode discrimination task. Square brackets indicate that the preceding electrode cannot be discriminated from all electrodes enclosed in the square brackets, e.g., 7-[9-11] indicates that electrode 7 is indiscriminable from electrodes 9, 10, and 11. The number of pairs of indiscriminable electrode pairs ranged from 1 to 11 across subjects. This range is similar to data reported in Zwolan *et al.* (1997). The number of indiscriminable electrode pairs has been shown to be correlated with these subjects' speech recognition abilities (more indiscriminable electrode pairs being indicative of poorer speech recognition abilities) in Throckmorton and Collins (1999). A correlation between electrode discrimination and speech recognition has also been reported by Henry *et al.* (1997).

B. Pitch-ranking task

Following the approach described in Collins *et al.* (1997), stimulus response matrices were calculated based on subjects' responses to all of the combinations of the stimuli presented during the final ten replications of the stimulus (ten presentations of each possible pair). A row-sum analysis, based on a statistical theory developed specifically for analyzing the results of paired comparison experiments (David, 1988), was used to calculate the percent of the time a particular electrode pair was considered to be higher in pitch than the other electrode pairs in the area (percent wins).

Figure 1 plots the pitch estimates in terms of percent wins obtained using the row-sum analysis for each of the six subjects who participated in this study. As would be expected based on the tonotopic structure of an unimpaired cochlea, the pitch tends to decrease as the electrode number increases, since electrode pair 1 is located closest to the base of the cochlea. Clearly, not all of the pitch functions are monotonic functions of the electrode number, and three of the subjects (S2, S3, and S4) have pitch functions which exhibit fairly strong nonmonotonicities. These functions are similar to the more regular pitch functions reported by Collins *et al.* (1997) and Nelson *et al.* (1995). For all of the subjects, pitch appears to change at approximately the same rate as a function of distance along the electrode array. However, since the data are from a ranking as opposed to a scaling task, it is not possible to conclude whether the changes in pitch from electrode pair to electrode pair are of the same magnitude.

Using the percent wins data that are calculated from the row-sum analysis, a least-significant difference analysis was performed to determine whether each electrode pair was significantly different in pitch when compared to the remaining

electrode pairs (David, 1988). Based on the results presented in Collins *et al.* (1997), a d' analysis, such as was suggested by Townshend *et al.* (1987) and Nelson *et al.* (1995), was not used since it requires the assumptions of (1) a Gaussian distribution of the perceptual variable; (2) a unidimensional percent; and (3) identical variance for each electrode (Green, 1970). The set of indiscriminable electrode pairs as determined by the analysis of the pitch-ranking data is listed in the third column of Table III for each subject. In all cases, the number of electrodes determined to be indiscriminable using this method exceeds the number of indiscriminable electrodes measured during the electrode discrimination experiment. Again, these results are consistent with those reported by Collins *et al.* (1997). In particular, they indicate that two electrodes may have the same pitch but still be discriminable (see, for example, pitch-ranking data for electrodes 12 and 13 for subject S1 (Fig. 1) and electrode discrimination data for electrodes 12 and 13 for subject S1 (second column, Table III). These results may suggest a second percept associated with the stimulation.

To further address this issue, the row-sum formulation can be used to estimate whether the evoked percept is unidimensional in addition to calculating discriminability (David, 1988). A measure of the consistency of the ranking data, which is related to the number of transitivity violations within each set of data, was calculated for each of the subjects. Thus, if in the pitch-ranking data electrode 1 is consistently ranked higher than 2, 2 higher than 3, and 3 higher than 1, this provides evidence that subjects cannot provide accurate ranking data based solely on a pitch percept. The consistency measure is also listed in Table III. A self-consistent data set would have consistency measure 1, whereas a completely random data set would have consistency measure of 0. The range of consistency measured across the subjects in this study was 0.567 to 0.737, which is higher on average than the values reported in Collins *et al.* (1997), but still within the range reported in that study. These values, like those reported previously, provide additional evidence that there may be multiple percepts associated with electrical stimulation at various places in the cochlea.

C. Multidimensional scaling task

Dissimilarity matrices were computed from the experimental data using the method described by Schiffman *et al.* (1981). The goal of an MDS analysis is to arrange the stimuli in an N-dimensional space so that the error between the distances that were calculated in the solution space and the distances that were measured experimentally is minimized. The error between the solution and the data is referred to as the ''stress,'' and an R^2 (RSQ) value can be computed which measures the proportion of the optimally scaled data that is accounted for by the model. Both metrics measure goodness of fit. In the solution space, two stimuli that were judged ''most different'' in the experiment will be far apart in the stimulus space.

In order to determine and then interpret a valid multidimensional solution for the data sets, it is necessary to first determine the number of dimensions that are required to represent the stimulus space accurately. There are two steps to this process, the first of which is measuring goodness of fit and the second of which is interpreting the dimensions. To achieve the first step, the MDS analysis was performed using an increasing number of dimensions, $N=1$, 2, 3, 4, and 5. For each solution, the stress and RSQ values were recorded. The location of a change in the slope of the stress or RSQ versus dimension plots indicates the appropriate number of dimensions for the solution (Schiffman *et al.*, 1981). The RSQ and stress results are plotted versus the number of dimensions used in the solution for each of the subjects in Figs. 2 and 3, respectively. The solutions were obtained us-

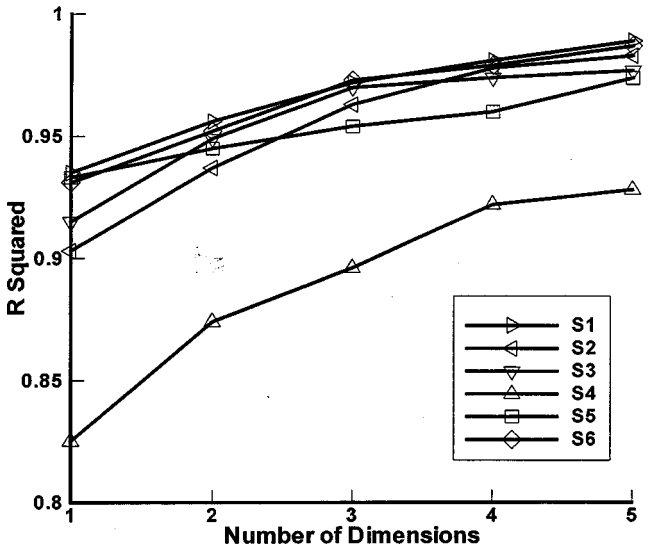


FIG. 2. R^2 (RSQ) value reported by the MDS algorithm plotted as a function of the number of dimensions in the MDS solution. The results are plotted for subject S1 with a right-pointing triangle, subject S2 with a left-pointing triangle, subject S3 with a downward triangle, subject S4 with an upward triangle, S5 with a square, and subject S6 with a diamond.

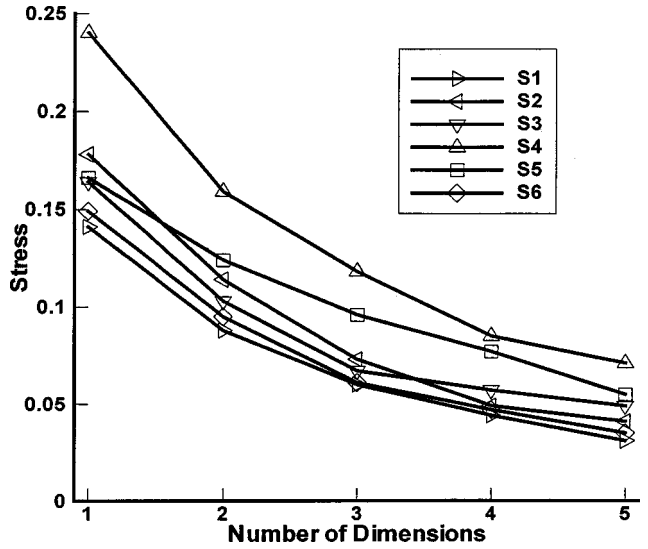


FIG. 3. Stress reported by the MDS algorithm plotted as a function of the number of dimensions in the MDS solution. The results are plotted for subject S1 with a right-pointing triangle, subject S2 with a left-pointing triangle, subject S3 with a downward triangle, subject S4 with an upward triangle, S5 with a square, and subject S6 with a diamond.

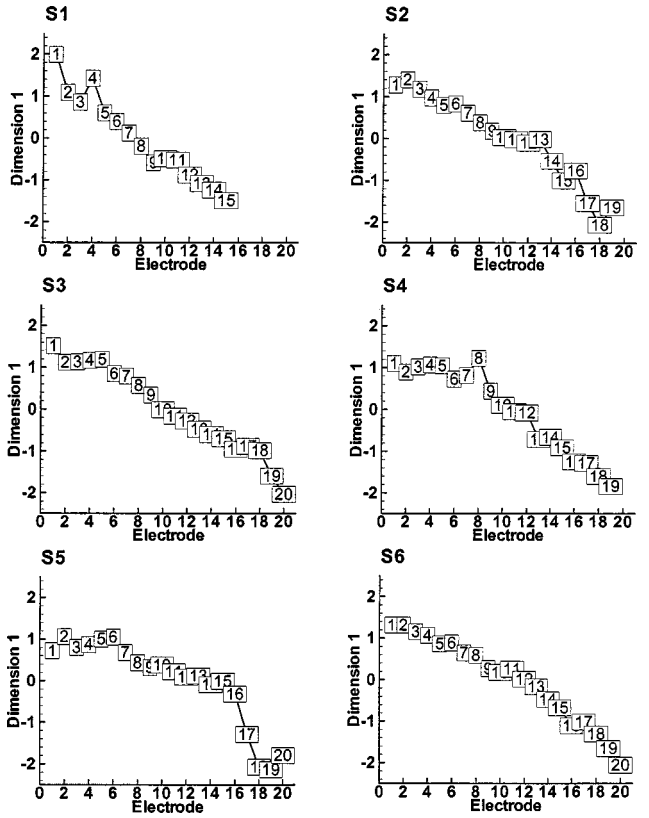


FIG. 4. The one-dimensional solution spaces obtained from the MDS analysis for each of the six subjects. The perceptual location of each electrode is plotted along the ordinate as a function of the electrode number (abscissa). A box enclosing the number of the basal member of the electrode pair denotes the location of a particular electrode. Lines attach the electrodes sequentially.

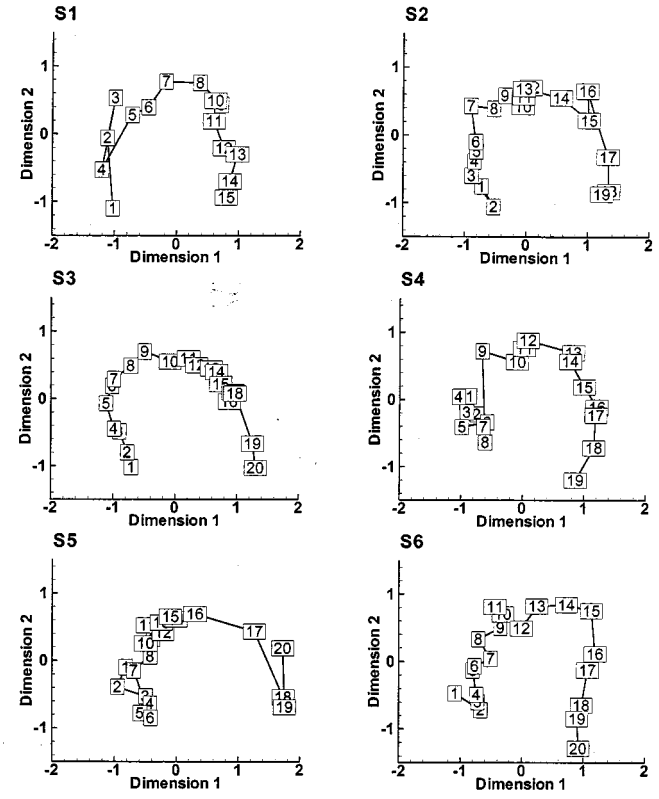


FIG. 5. The two-dimensional solution spaces obtained from the MDS analysis for each of the six subjects. The perceptual location of each stimulus is plotted in dimension 1 along the abscissa and dimension 2 along the ordinate. A box enclosing the number of the basal member of the electrode pair denotes the location of a particular electrode. Lines attach the electrodes sequentially.

ing dissimilarity matrices calculated from the data averaged across all blocks for each subject. Although no clear change in slope is observed in either the RSQ or stress curves, it is clear that there is a substantial drop in the stress and increase in the RSQ as the number of dimensions increases from one to two, and in some cases from two to three. This lack of a knee in the curve is similar to the results described by McKay *et al.* (1996) for their dual-electrode stimulation. The RSQ values reported here for the one- and two-dimensional solutions are of the same order of magnitude as their two- and three-dimensional solutions. No substantial changes in slope are observed when increasing the number of dimensions beyond three for any subject. In the remainder of the analysis, the solutions obtained via the one- and two-dimensional analyses are considered.

The one-dimensional MDS solutions generated by SYSTAT for each of the subjects are shown in Fig. 4. The location of each electrode in the solution space (ordinate) is plotted as a function of electrode number (abscissa), and the locations are also labeled with the electrode number for ease of interpretation. The one-dimensional solution space is an ordered function of the electrode number, with the location decreasing as electrode number increases. These results indicate that if a one-dimensional solution completely captures the structure of the perceptual space, the percept associated with place of stimulation follows a more-or-less orderly progression from base to apex. Given this fact, and by comparing the results presented in Fig. 4 with those contained in Fig. 1, it appears that, as expected, there is a relationship between the percept mapped by the one-dimensional solution and the pitch percept. The patterns of the results are similar for each of the subjects, although there are some differences. For example, Subject S2 has a strong nonmonotonicity in his pitch structure between electrodes 11 and 15 which is not as clearly evident in the MDS results in Fig. 4. The relationship between the MDS solution and the pitch data will be explored more fully in the next section. There are also regions for certain subjects (e.g., electrodes 1–5 for subject S4) where the percept does not appear to be changing—i.e., the electrodes overlap in the solution space. It is possible that such regions may indicate indiscriminable electrodes—an analysis that will also be pursued further in the next section.

The two-dimensional MDS solution generated by SYSTAT provides a location in two-dimensional space for the percept corresponding to stimulation of each electrode pair. Note that these dimensions are not guaranteed to be orthogonal. Nonetheless, the ''location'' of each electrode can be plotted on a standard $x-y$ plot, where each axis corresponds to one of the two dimensions in the solution space. Figure 5 provides the MDS solutions for the data obtained from each of the subjects. The abscissa corresponds to the location of the perceptual response to stimulation of an electrode pair in ''dimension 1,'' and the ordinate corresponds to the location

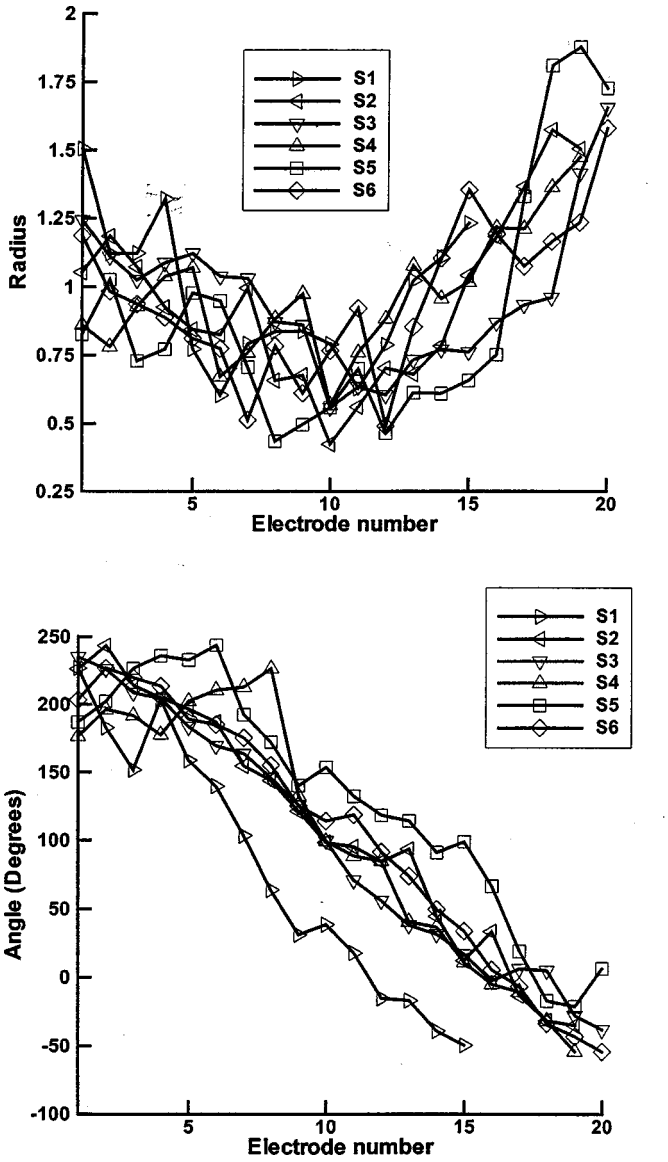


FIG. 6. Plots of the radius (top panel) and angle (bottom panel) calculated from the polar formulation of the two-dimensional MDS solutions as a function of electrode number. The results are plotted for subject S1 with a right-pointing triangle, subject S2 with a left-pointing triangle, subject S3 with a downward triangle, subject S4 with an upward triangle, S5 with a square, and subject S6 with a diamond.

of the electrode pair in ''dimension 2.'' The electrode numbers corresponding to each location are indicated. For ease of interpretation, lines have been drawn between adjacent electrode pairs.

There are several patterns that are common to the two-dimensional MDS solutions for all subjects. First, the values taken on by each electrode pair in the first dimension tend to be a monotonic function of the electrode number. This result, coupled with the results obtained on the pitch-ranking task, may indicate that the first dimension of the two-dimensional solution can be interpreted as the pitch. This result is similar to that obtained from the one-dimensional solution, and will be explored further in the next section.

Second, the pattern formed by connecting adjacent electrode pairs by lines in the two-dimensional solution space is ''U''-shaped for all subjects. A similar U-shaped pattern was observed for the single-electrode data tested by McKay *et al.* (1996). However, their stimulus set consisted of both single- and dual-electrode stimuli so the two-dimensional nature of that stimulus set can only be concluded for the set as a whole. If the first dimension in the solution space is interpreted as pitch, the U-shaped pattern could indicate that a second, nonpitch, dimension associated with electrical stimulation changes little for more centrally located electrode pairs and more substantially for electrode pairs located near the base or the apex.

As suggested by McKay *et al.* (1996) in their analysis of the U-shaped pattern of their two-electrode stimuli, since the values taken on by each electrode pair are distributed more or less sequentially around the circumference of the U, it is possible that the angle associated with the location of a particular electrode pair could be interpreted as the pitch associated with stimulation of an electrode pair. The validity of this assumption is pursued in the next section. Figure 6 plots the values corresponding to the radial and angular dimension for each electrode pair obtained from the polar formulation of the two-dimensional solution for each subject. Clearly, a U-shaped pattern is still present in the radial dimension. These data could also indicate a percept that is more salient at the ends of the electrode array than near the center.

For most subjects, there are sets of electrodes whose two-dimensional coordinates are very close in the solution space, thus causing them to appear to ''overlap.'' It is possible that electrodes whose locations in the MDS solution space overlap are indiscriminable. This hypothesis is also investigated further in the next section.

TABLE IV. Pearson correlation coefficient calculated between subject's pitch data and multidimensional scaling data. Each column lists the number of dimensions, *N*, in the total solution as well as the dimension being correlated with the pitch data. Bold entries are significant ($p<0.001$).

Subject	$N=1$ Dimension 1	$N=2$ Dimension 1	$N=2$ Dimension 2
S1	**0.967**	**0.972**	−0.085
S2	**0.898**	**0.866**	−0.361
S3	**0.928**	**0.857**	−0.255
S4	**0.964**	**0.919**	−0.001
S5	**0.898**	**0.844**	−0.444
S6	**0.981**	**0.899**	0.059

D. Analysis of the perceptual dimensions

In order to determine whether the various dimensions associated with a particular MDS solution space correspond to the psychophysical attribute of pitch, correlation analyses were performed. The values associated with each electrode pair in each dimension of the one- and two-dimensional MDS solutions were correlated with the pitch data obtained from each subject. The correlation coefficients obtained from a Pearson correlation analysis are listed in Table IV. The Pearson analysis assumes that the data follow a Gaussian distribution. A Spearman correlation analysis was also performed in order to avoid assumptions regarding the probability density functions associated with the underlying data. The correlation coefficients for the Spearman correlation analysis

TABLE V. Spearman correlation coefficient calculated between subject's pitch data and multidimensional scaling data. Each column lists the number of dimensions, *N*, in the total solution as well as the dimension being correlated with the pitch data. Bold entries are significant ($p<0.001$).

Subject	$N=1$ Dimension 1	$N=2$ Dimension 1	$N=2$ Dimension 2
S1	**0.936**	**0.918**	0.046
S2	**0.919**	**0.837**	−0.251
S3	**0.934**	**0.866**	−0.162
S4	**0.892**	**0.820**	0.047
S5	**0.886**	**0.836**	−0.564
S6	**0.992**	**0.901**	−0.039

TABLE VI. Pearson correlation coefficient calculated between subject's pitch and electrode data for dimension 1 of the 1D and 2D solutions and the comparison to radius and angle. Bold entries are significant ($p<0.001$).

Subject	$N=1$ Dimension 1	$N=2$ Dimension 1	$N=2$ Radius	$N=2$ Angle
S1	**0.967**	**0.972**	0.441	**0.970**
S2	**0.898**	**0.866**	−0.184	**0.920**
S3	**0.928**	**0.857**	0.047	**0.931**
S4	**0.964**	**0.919**	−0.723	**0.974**
S5	**0.898**	**0.844**	−0.492	**0.937**
S6	**0.981**	**0.899**	−0.522	**0.979**

are listed in Table V. Any correlation implies that the data are linearly related to the percept used for interpretation.

The correlation coefficients presented in Tables IV and V indicate that the pitch data are strongly correlated with the first dimension of both the 1D and 2D solutions spaces, and that the correlation coefficients are statistically significant. The pitch data are not significantly correlated with the second dimension of the 2D solution space. The Pearson correlation coefficients are largest for the first dimension of the 1D solution for four of the subjects, and are largest for the first dimension of the 2D (S1) solution for one subject. The Spearman correlation coefficients are largest for the first dimension of the 1D solution for all subjects. In all cases, the differences between the correlation coefficients between the pitch data and the first dimension of the solution space for the 1- and 2D solutions are not large. This result is consistent with the hypothesis that the first dimension of the MDS solution space can be interpreted as pitch.

Although the first dimension of the MDS solution space may be interpreted as pitch, the ''circular'' form of the solution space and the orderly structure therein suggests that it is also possible that the angle around the circle may be interpreted as pitch [this was also considered for dual-electrode stimuli by McKay *et al.* (1996)]. To investigate this hypothesis, the two-dimensional MDS solutions can be represented in polar coordinates instead of the rectangular coordinates used in Fig. 5. In order to determine whether either of the two dimensions associated with a polar solution is more strongly correlated with the pitch data than the rectangular coordinates, the rectangular solutions were converted to polar coordinates. (Figure 6 plots the radius and angle estimates as a function of electrode number for each of the six subjects.) The data from the two polar dimensions, radius and angle, associated with each subject were then correlated with the pitch data. The results of a Pearson and Spearman correlation analysis on this data are presented in Tables VI and VII, respectively. The correlation data indicate that, as was hypothesized, the correlation between the pitch data and the angle dimension of the two-dimensional MDS solution is statistically significant for all subjects. Furthermore, these correlation coefficients are higher than the correlation between the pitch data and the first dimension of the other two solutions for four of the six subjects. However, this is simply a trend; the differences between the correlation coefficients are small, and the sample size is too small to test significance of these differences. Thus, the differences must be viewed as not significant, and no firm conclusions can be drawn.

The MDS and pitch-ranking analyses indicate that there may be multiple percepts associated with electrical stimulation across the cochlea, since the goodness-of-fit measures associated with the MDS data improve with the number of dimensions and the consistency metric for the pitch data is substantially less than unity. Inconsistencies between the electrode discrimination data and the pitch-ranking analyses also provide evidence for multiple percepts. The correlation analyses described above also indicate that the first dimension of the two-dimensional solution is significantly correlated with the pitch data. Furthermore, the angular dimension of a polar formulation of the two-dimensional solution is strongly correlated with the pitch data. Although loudness balancing was performed prior to gathering the MDS data, it is possible that one of the additional dimensions is an artifact due to improper loudness balancing. However, it is difficult to believe such an artifact would be so consistent across subjects. The loudness-balance procedure that was used, referred to as ''reference'' balancing, was one in which the subject balanced the loudness of all electrode pairs to the loudness of a centrally located reference electrode pair. It is conceivable that if all subjects' loudness balances were subject to errors which increased monotonically as a function of the distance from the reference electrode pair (and in the same perceptual direction—i.e., either louder or softer) that the MDS solutions would show such patterns.

In order to investigate this hypothesis, an alternative loudness-balance procedure, referred to as ''adjacent'' balancing, was performed by one of the six subjects, and a second set of MDS data was gathered and analyzed using the new loudness-balanced values. In the second loudness-balance procedure, the most basal electrode pair was initially set to a comfortable loudness level, and then the electrode

TABLE VII. Spearman correlation coefficient calculated between subject's pitch and electrode data for dimension 1 of the 1D and 2D solutions and the comparison to radius and angle. Bold entries are significant ($p<0.001$).

Subject	$N=1$ Dimension 1	$N=2$ Dimension 1	$N=2$ Radius	$N=2$ Angle
S1	**0.936**	**0.918**	0.236	**0.943**
S2	**0.919**	**0.837**	−0.067	**0.897**
S3	**0.934**	**0.866**	0.241	**0.944**
S4	**0.892**	**0.820**	−0.691	**0.906**
S5	**0.886**	**0.836**	0.006	**0.875**
S6	**0.992**	**0.901**	−0.451	**0.989**

TABLE VIII. Significance level of *t*-test between two sets of distances for each subject across the 1D and 2D MDS solutions (NS=not significant).

Subject	Prob(1D)	Prob(2D)
S1	0.05	0.1
S2	NS	NS
S3	0.01	0.1
S4	NS	NS
S5	NS	NS
S6	0.01	0.01

pair adjacent to the reference electrode was balanced using six successive adjustments (exactly as was done for the original loudness-balance procedure). Next, the ''adjusted'' electrode was used as the reference electrode (with its adjusted loudness-balance level), and the next most apical electrode pair was balanced to this new reference electrode. This process continued until all electrode pairs had been balanced. This procedure has been described more completely in Zwolan *et al.* (1997).

The results of this analysis indicated no substantial differences in the one- or two-dimensional solutions evaluated from the data obtained under different loudness-balancing procedures. This result lends credence to the claim that the characteristics associated with the additional percept measured using the MDS procedure are not related to variability in the loudness of the stimuli. Thus, the fact that the solutions for both methods are very similar suggests that any loudness differences between stimuli were not a factor influencing the location of the points in space.

E. Discrimination analysis

The final analyses were performed in an attempt to relate the Euclidean distance between electrode pairs in MDS space to perceptual distance. First, a *t*-test was used to compare the average distance between electrode pairs found to be indiscriminable in the electrode discrimination task to the average distance between electrodes found to be discriminable. The analysis was performed for each subject for the one- and two-dimensional solutions, and the results are summarized in Table VIII. Although the difference was statistically significant for some subjects, distance was not a clear predictor of indiscriminable electrodes in this analysis.

The goal of the final analysis was to determine whether indiscriminable electrodes can be more accurately pinpointed using an MDS analysis than a pitch-ranking analysis. For the MDS data, the results were sorted into distances associated with discriminable electrodes, and distances associated with indiscriminable electrodes for each subject. Based on the set of distances associated with indiscriminable electrodes, a ''threshold'' was determined for each subject that would guarantee selection of 70% of the indiscriminable electrodes. This threshold was then applied to the set of discriminable electrodes in order to determine the number of ''false alarm'' electrodes, i.e., the number of discriminable electrodes that would falsely be determined to be indiscriminable based on the 70% threshold. The results for each subject are listed in Table IX, both in terms of the number of discriminable electrodes selected and the associated probability of false alarm. A similar analysis was performed using the pitch-ranking data. A threshold was determined for each subject to guarantee selection of 70.7% of the indiscriminable electrodes based on a difference in the percent wins scores. Table IX also lists the number of discriminable electrodes selected using this threshold, and the associated probability of false alarm. For all subjects but one, the number of discriminable electrodes falsely selected was lower using the MDS data than the pitch-ranking data. If, as some studies indicate [Zwolan *et al.* (1997); Henry *et al.* (1997); Throckmorton and Collins (1999)], electrode discrimination measures are clinically relevant, finding accurate and efficient psychophysical measures of electrode discrimination that are efficient should be of interest to the clinical community.

IV. DISCUSSION

Previous studies investigating the percepts evoked by individual stimulation of each electrode pair in the array of implanted subjects have suggested that pitch is generally consistent with tonotopic order. Whether or not there are additional characteristics associated with the percept has been partially addressed in previous studies. The goals of this study were to investigate the hypothesis that percepts other than pitch associated with electrode place may be elicited, to investigate the structure of the perceptual dimensions, and to determine whether the structure of the perceptual space could be linked to specific psychophysical

TABLE IX. Probability of false alarm and associated number of electrodes exceeding a threshold for the MDS (distance) and pitch ranking (difference in percent wins) data. Threshold was calculated to obtain a probability of detection of 70%.

Subject	**MDS** Probability of false alarm using threshold for probability of detection=70	**MDS** Number of false alarm electrodes for probability of detection=70	**Pitch ranking** Probability of false alarm using threshold for probability of detection=70	**Pitch ranking** Number of false alarm electrodes for probability of detection=70
S1	7	7	19	19
S2	13	22	14	23
S3	6	11	12	22
S4	6	10	20	34
S5	12	22	7	13
S6	2	3	4	8

measures, such as pitch or electrode discriminability. To achieve this goal, the loudness of stimuli presented to each electrode pair was balanced, and six subjects performed pitch ranking, electrode discrimination, and multidimensional scaling tasks.

Results from all tasks were consistent with the results presented for the individual tasks in previous studies. In addition, the pattern of results was fairly consistent across subjects. Specifically, analysis of the pitch-ranking data failed to accurately predict indiscriminable electrodes, and always predicted more indiscriminable electrodes than the electrode discrimination task. Analysis of the transitivity in the stimulus ordering from the pitch-ranking data indicated the possibility that the percepts were multidimensional in nature, a fact which, as Shannon (1983b) noted, may render pitch evaluation more difficult, and thus the discriminability estimates more prone to error. The MDS results also indicated that a multidimensional percept might be present. The data also indicated that the first dimension of a two-dimensional solution could be interpreted as pitch, and that the second dimension exhibited a characteristic curvature as a function of the first dimension.

An important step in any MDS analysis is to interpret the dimensions of the solution space. Strong correlations were observed between the values taken on by the electrode pairs in the first dimension of each of the Cartesian MDS solutions and the pitch associated with each electrode pair. These strong correlations may indicate that the first dimension in the perceptual space can be interpreted as being linearly related to the pitch of the stimulus. The two-dimensional solutions were also formulated in polar coordinates, and a strong correlation was observed between the angle associated with each electrode pair and its pitch value. The curvature of the plots obtained from the rectangular formulation of the MDS solution suggests that the second dimension varies more quickly at the ends of the array than in the center. The U-shaped pattern observed in Fig. 5 may also indicate a percept that changes more quickly at the ends of the electrode array than near the center. Since the array is normally inserted approximately $1\frac{1}{4}$ turns of the cochlea, currents generated by the most basal and apical electrodes may stimulate partially overlapping populations of neurons. Such overlap could be one source of the perceptual similarity in the second dimension for stimuli generated by the basal- and apical-most electrodes. An alternative explanation may be that the array is closer to the underlying neurons in the center of the array, and thus that additional current spread near the ends of the array results in a second percept in the response.

One hypothesis to explain the basis of the curvature is that the *rate* of the stimulus as a function of electrode place generated a difference in the percept across the electrode array, and it is this phenomenon which is captured in the second perceptual dimension. Although the rate was kept constant as a function of electrode place, it is possible that the rate–pitch percept and the place–pitch percept interacted differently as a function of place, and that this interaction could have generated a timbre-like percept (if the rate does not interact differentially no additional percept will be created). A natural hypothesis, based on the natural frequency coding of the cochlea and the depth of insertion of the electrode array, would be that a monotonic (or linear) relationship would exist between the two perceptual dimensions. Such a relationship could occur since all the electrodes should have a ''natural'' pitch that is higher than that evoked by a 500-pps rate. However, the data obtained in this study indicate a U-shaped or curved, as opposed to monotonic, relationship. This discrepancy can be addressed by considering the nature of the data obtained in the MDS task.

In the MDS task, subjects were asked to provide an estimate of the difference, or dissimilarity, between all possible pairs of stimuli. Prior to each trial, all of the stimuli were presented to the subjects so that they knew the perceptual extent of the space to be explored. In an ideal situation, only the most ''distant'' stimuli would be given a distance measure of ''1.'' However, for a given electrode, subjects would often ascribe a distance of 1 to a range of electrodes. For example, when judging the perceptual distance between all possible electrodes and electrode 1, subjects often assigned a distance of 1 to electrodes 18, 19, and 20. It is difficult to determine whether these distance judgments were accurate, i.e., electrode 1 was equally far perceptually from electrodes 18, 19, and 20, or whether subjects effectively apply a ''limiter,'' i.e., are incapable of judging perceptual distance beyond some maximum distance.

If, in fact, subjects are limited in their ability to judge perceptual distances, this limitation could affect the results of the MDS analysis by ''curving'' an otherwise ''straight'' two-dimensional perceptual space. For example, consider a perceptual space where 11 stimuli are equally spaced in two perceptual dimensions, are ordered in the first perceptual dimension, and the relationship between the two perceptual dimensions is linear. An MDS solution for a noise-free estimate of this perceptual space is shown in the top panel of Fig. 7. If a subject is able to provide noise-free estimates of this perceptual space, but is limited in their ability to provide distance estimates beyond some maximum level, the effects can be modeled by modifying the hypothesized dissimilarity matrix. In the example described above, all stimuli further than 6 units apart were assumed to be equally dissimilar and given a distance measure of 1. This implies, for example, that stimulus 1 is equally far from stimuli 8–11.

The MDS solution in three dimensions for this problem is shown in the bottom panels of Fig. 7. Clearly, this inability to define distance in an otherwise two-dimensional linear perceptual space results in an artificial curvature of the dimension 1–dimension 2 space, an oscillatory behavior in the dimension 1–dimension 3 space (which should be zero), and a looping behavior in the dimension 2–dimension 3 space. In addition, the two-dimensional space requires three dimensions to be adequately represented, the third of which is not perceptually based. The curvature results because stimulus 1 is required to be equally distant from stimuli 8–11, stimulus 2 is required to be equally distant from stimuli 9–11, etc. The best solution can be visualized as placing the middle points along an arc on a sphere. The ends of the arc then rotate perpendicularly to each other so as to minimize the error across all of the distances. The one-dimensional solu-

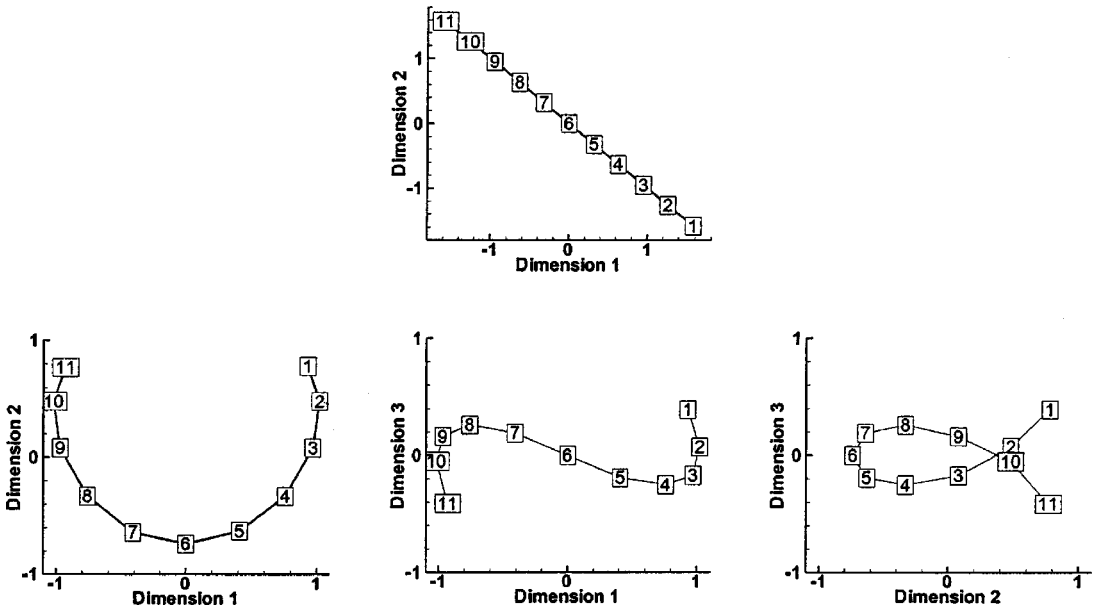


FIG. 7. Top panel: Perceptual space for the example in which 11 stimuli are equally spaced in two perceptual dimensions. Bottom panel: Three-dimensional MDS solution space where all stimuli further than 6 units apart are assumed to be equally dissimilar. This simulates the case where a subject has limited ability to resolve distances greater than some particular perceptual distance. In the leftmost plot, the perceptual location of each stimulus is plotted in dimension 1 along the abscissa and dimension 2 along the ordinate. In the center plot, the perceptual location of each stimulus is plotted in dimension 1 along the abscissa and dimension 3 along the ordinate. In the rightmost plot, the perceptual location of each stimulus is plotted in dimension 2 along the abscissa and dimension 3 along the ordinate.

tion to this problem is similar to the top panel of Fig. 7 (stress=0.11); the two-dimensional solution resembles the left panel of the bottom portion of Fig. 7 (stress=0.01), and the three-dimensional solution has a stress of 0.002.

The data obtained from all of the subjects who participated in this study are consistent with the pattern predicted in the bottom panels of Fig. 7. In addition, all subjects would ascribe a distance of ''1'' to a range of apical electrodes for a given basal electrode. Thus, it is possible that subjects are unable to provide accurate distance measures once stimuli are perceptually distant, and the underlying percept is actually linear (or monotonic) in a two-dimensional perceptual space. Note that the example provided in Fig. 7 does not preclude a one-dimensional solution; however, additional lines of evidence suggest at least a two-dimensional solution.

This analysis could support the hypothesis that the second perceptual dimension is related to an interaction between rate–pitch and place–pitch phenomena if the assumption that the perceptual response to the interaction would change monotonically with place. In addition, based on the results shown in Fig. 7, it is clear that a projection of the curved MDS solution onto dimension 1 should not provide as accurate a prediction of the pitch data as does the angular coordinate of a polar representation. The relatively stronger correlation between the angle dimensions in the polar coordinate data and the pitch dimension follows this pattern, and thus may support the hypothesis that the second perceptual dimension is a monotonic function of the first, as opposed to a curved function.

A concern in attempting to make conclusions regarding the dimensionality of the perceptual response results from a ''horseshoe effect'' that has been observed in the literature (Kendall, 1971). In general, if a *single-dimensional* set of stimuli is subject to a two-dimensional MDS analysis, patterns similar to a horseshoe (and similar to those observed in this data set) are obtained. In fact, similar horseshoe effects have been recorded for MDS applied to rate of stimulation data (McKay and Carlyon, 1999). In this case, the authors concluded that the single arc in their two-dimensional space could not be taken as an indication that there are two separate dimensions because the arrangement of the points along the arc was similar to that in the one-dimensional space.

A final hypothesis that may be considered regarding this data is that while percepts other than pitch are elicited when stimulating electrodes along the array, there may not be reason to assume that these percepts will be ordered along a *single additional dimension* or that they would have a monotonic structure (which is required by the MDS analysis technique). For example, one end of the array may sound ''rougher'' while another end of the array may sound ''deeper.'' The data and analyses presented in this paper do not unequivocally indicate that the percepts are ordered along a single dimension, although the confluence of evidence from the multiple tasks does appear to indicate that in some cases there is more than one percept associated with electrode place. In this case, it may be that the MDS solution for a single dimension (either the single-dimension MDS solution or the radial dimension of the two-dimensional solution) is more valid than the two-dimensional solution, or an MDS solution many not be an appropriate mechanism by which to investigate the structure of the perceptual space.

The final goal of this work was to investigate whether the MDS data could be used to accurately predict electrode discrimination data. Based on a threshold set to achieve ''detection'' of 70% of the indiscriminable electrodes, utilization of the distances obtained from the MDS solutions resulted in

fewer false alarms than utilization of the pitch differences obtained from the pitch-ranking task. This result indicates that the MDS data capture more of the relevant structure associated with the perceptual response to electrical stimulation, and also supports the hypothesis that the perceptual space is multidimensional.

Previous studies investigating the percept associated with electrode place have focused on the pitch-like dimension of the percept, and have not generally investigated alternate perceptual dimensions. The results of this study indicate that it may be possible to more effectively use electrode place for encoding speech information to improve both intelligibility as well as quality. For example, speech understanding may be improved by utilizing only those electrodes without substantial components within the non-place–pitch dimension, or these electrodes may be used to provide information regarding subtle temporal differences between different stimuli. Alternatively, if the rate–pitch and place–pitch percepts and their inter-relationships can be understood more clearly, a wider range of signals may be able to be represented with higher fidelity. Some work has already considered these issues (Tong *et al.*, 1983b; Pijl, 1997). Further work is necessary to investigate the perceptual structure and relevance of the additional dimensions and the percepts associated with them.

ACKNOWLEDGMENTS

We express our thanks and appreciation to the six subjects who gave us their time and effort, as well as Dr. C. McKay and Dr. R. Carlyon and the two anonymous reviewers for their valuable suggestions. Also, we thank Ms. Pat Roush, Dr. Debora Tucci, and Dr. Terry Zwolan for their help in contacting and recruiting subjects. In addition, we would like to express our appreciation to Dr. David Smith and Dr. Roger Miller for their input at various stages of this work. We are indebted to Cochlear Corporation for providing equipment necessary to carry out this study. This research was supported by NIH Grant No. 5R03-DC-03136-02.

Busby, P. A., Whitford, L. A., Blamey, P. J., Richardson, L. M., and Clark, G. M. (**1994**). ''Pitch perception for different modes of stimulation using the cochlear multiple-electrode prosthesis,'' J. Acoust. Soc. Am. **95**, 2658–2669.

Chatterjee, M., and Sharmon, R. V. (**1998**). ''Forward masked excitation patterns in multielectrode electrical stimulation,'' J. Acoust. Soc. Am. **103**, 2565–2572.

Collins, L. M., Zwolan, T. A., and Wakefield, G. H. (**1997**). ''Comparison of electrode discrimination, pitch ranking, and pitch scaling data in postlingually deafened adult cochlear implant subjects,'' J. Acoust. Soc. Am. **101**, 440–455.

David, H. A. (**1988**). *The Method of Paired Comparisons* (Oxford University Press, New York).

Dorman, M. F., Soli, S., Dankowski, K., Smith, L. M., McCandless, G., and Parkin, J. (**1990**). ''Acoustic cues for consonant identification by patients who use the Ineraid cochlear implant,'' J. Acoust. Soc. Am. **88**, 2074-2079.

Dowell, R. C., Martin, L. F. A., Tong, Y. C., Clark, G. M., Seligman, P. M., and Patrick, J. F. (**1982**). ''A 12-consonant confusion study on a multiple-channel cochlear implant patient,'' J. Speech Hear. Res. **25**, 509–516.

Eddington, D. K., Dobelle, W. H., Brackmann, D. E., Mladejovsky, M. G., and Parkin, J. (**1978a**). ''Auditory prostheses research with multiple channel intracochlear stimulation in man,'' Ann. Otol. Rhinol. Laryngol. **87** (Suppl. 53), 1–39.

Eddington, D. K., Dobelle, W. H., Brackman, D. E., Mladejovsky, M. G., and Parkin, J. (**1978b**). ''Place and periodicity pitch by stimulation of multiple scala tympani electrodes in deaf volunteers,'' Trans. Am. Soc. Artif. Intern. Organs **XXIV**, 1–5.

Eddington, D. K. (**1980**). ''Speech discrimination in deaf subjects with cochlear implants,'' J. Acoust. Soc. Am. **68**, 885–891.

Green, D. M. (**1970**). ''Application of detection theory in psychophysics,'' IEEE Proc. 713–723.

Henry, B. A., McKay, C. M., McDermott H. J., and Clark G. M. (**1997**). ''The relationship between speech information perceived by cochlear implantees in different spectral regions and electrode discrimination,'' 1997 Conference on Implantable Auditory Prostheses, Asilomar Conference Center, Pacific Grove, California.

Kendall, D. G. (**1971**). ''Seriation from abundance matrices,'' in *Mathematics in Archeological and Historical Sciences*, edited by F. R. Hodson, D. G. Kendall, and P. Tautu (Edinburgh University Press, Edinburgh).

McKay, C. M., McDermott, H. J., and Clark, G. M. (**1996**). ''The perceptual dimensions of single-electrode and nonsimultaneous dual-electrode stimuli in cochlear implantees,'' J. Acoust. Soc. Am. **99**, 1079–1090.

McKay, C. M., and Carlyon, R. C. (**1999**). ''Dual temporal pitch percepts from acoustic and electric amplitude-modulated pulse trains,'' J. Acoust. Soc. Am. **105**, 347–357.

Nelson, D. A., Van Tasell, D. J., Schroder, A. C., Soli, S., and Levine, S. (**1995**). ''Electrode ranking of 'place pitch' and speech recognition in electrical hearing,'' J. Acoust. Soc. Am. **98**, 1987–1999.

Pauka, C. K. (**1989**). ''Place–pitch and vowel–pitch comparisons in cochlear implant patients using the Melbourne-Nucleus cochlear implant,'' J. Laryngol. Otol. Suppl. **19**, 1–31.

Pfingst, B. E. (**1988**). ''Comparisons of psychophysical and neurophysiological studies of cochlear implants,'' Hear. Res. **34**, 243–252.

Pijl, S. (**1997**). ''Pulse rate matching by cochlear implant patients: effects of loudness randomization and electrode position,'' Ear Hear. **18**, 316–325.

Schiffman, S. S., Reynolds, M. L., and Young, F. W. (**1981**). *An Introduction to Multidimensional Scaling: Theory, Methods, and Applications* (Academic, New York).

Shannon, R. V. (**1983a**). ''Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics,'' Hear. Res. **11**, 157–189.

Shannon, R. V. (**1983b**). ''Multichannel electrical stimulation of the auditory nerve in man. II. Channel interaction,'' Hear. Res. **12**, 1–16.

Simmons, F. B. (**1966**). ''Electrical stimulation of the auditory nerve in man,'' Arch. Otolaryngol. **84**, 2–54.

Soli, S. (**1990**). ''Perceptual evaluation of a neurally based encoding strategy for cochlear implants,'' in *Cochlear Implants*, edited by J. F. Miller and F. A. Spelman (Springer, New York).

Throckmorton, C. S., and Collins, L. M. (**1999**). ''Investigation of the effects of temporal and spatial interactions on speech recognition skills in cochlear implant subjects,'' J. Acoust. Soc. Am. **105**, 861–873.

Tong, Y. C., Millar, J. B., Clark, G. M., Martin, L. F., Busby, P. A., and Patrick, J. F. (**1980**). ''Psychophysical and speech perception studies on two multiple channel cochlear implant patients,'' J. Laryngol. Otol. **94**, 1241–1256.

Tong, Y. C., Clark, G. M., Blamey, P. J., Busby, P. A., and Dowell, R. C. (**1982**). ''Psychophysical studies for two multiple-channel cochlear implant patients,'' J. Acoust. Soc. Am. **71**, 153–160.

Tong, Y. C., Dowell, R. C., Blamey, P. J., and Clark, G. M. (**1983a**). ''Two-component hearing sensations produced by two-electrode stimulation in the cochlea of a deaf patient,'' Science **219**, 993–994.

Tong, Y. C., Blamey, P. J., Dowell, R. C., and Clark, G. M. (**1983b**). ''Psychophysical studies evaluating the feasibility of a speech processing strategy for a multiple-channel cochlear implant,'' J. Acoust. Soc. Am. **74**, 73–80.

Tong, Y. C., and Clark, G. M. (**1985**). ''Absolute identification of electric pulse rates and electrode positions by cochlear implant patients,'' J. Acoust. Soc. Am. **77**, 1881–1888.

Townshend, B., Cotter, N., Van Compernolle, D., and White, R. L. (**1987**). ''Pitch perception by cochlear implant subjects,'' J. Acoust. Soc. Am. **82**, 106–115.

Zwolan, T. A., Collins, L. M., and Wakefield, G. H. (**1997**). ''Electrode discrimination and speech perception in postlingually deafened adult cochlear implant subjects,'' J. Acoust. Soc. Am. **102**, 3673–3685.

Evaluation of feedback-reduction algorithms for hearing aids

Julie E. Greenberg, Patrick M. Zurek,[a] and Merry Brantley[b]
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 17 April 2000; revised 28 June 2000; accepted 15 August 2000)

Three adaptive feedback-reduction algorithms were implemented in a laboratory-based digital hearing aid system and evaluated with dynamic feedback paths and hearing-impaired subjects. The evaluation included measurements of maximum stable gain and subjective quality ratings. The continuously adapting CNN algorithm (Closed-loop processing with No probe Noise) provided the best performance: 8.5 dB of added stable gain (ASG) relative to a reference algorithm averaged over all subjects, ears, and vent conditions. Two intermittently adapting algorithms, ONO (Open-loop with Noise when Oscillation detected) and ONQ (Open-loop with Noise when Quiet detected), provided an average of 5 dB of ASG. Subjects with more severe hearing losses received greater benefits: 13 dB average ASG for the CNN algorithm and 7–8 dB average ASG for the ONO and ONQ algorithms. These values are conservative estimates of ASG because the fitting procedure produced a frequency-gain characteristic that already included precautions against feedback. Speech quality ratings showed no substantial algorithm effect on pleasantness or intelligibility, although subjects informally expressed strong objections to the probe noise used by the ONO and ONQ algorithms. This objection was not reflected in the speech quality ratings because of limitations of the experimental procedure. The results clearly indicate that the CNN algorithm is the most promising choice for adaptive feedback reduction in hearing aids. [S0001-4966(00)04111-4]

PACS numbers: 43.66.Ts, 43.60.Lq [SPB]

I. INTRODUCTION

Acoustic feedback is a common problem in hearing aids, limiting the maximum gain available to the user. When this maximum gain is exceeded, the hearing aid signal oscillates, producing an annoying sound referred to as ''whistling,'' ''screeching,'' or ''howling.''

The essence of the problem is shown in Fig. 1, which illustrates a generic digital hearing aid system. The signal received by the microphone consists of two components, the desired acoustic input and an undesired feedback signal related to the hearing aid output. The stability of the system is governed by the loop gain, which is comprised of the product of the frequency responses of all of the components in the feedforward and feedback paths, that is the preamplifier, the A/D converter, the digital hearing aid system, the D/A converter, and the acoustic feedback path. Whenever the magnitude of this product exceeds unity, the system is unstable and prone to oscillation.

Previous work on this problem has considered a number of approaches to reducing the effects of acoustic feedback in order to permit the hearing aid user to obtain higher stable gain. The most successful approaches to date consist of adaptively estimating and canceling the acoustic feedback path. These adaptive feedback-reduction algorithms can be thought of as belonging to one of two classes: continuously adapting algorithms, in which the adaptive filter weights are adjusted during normal closed-loop operation of the hearing aid; and noncontinuously adapting or open-loop algorithms, in which the hearing aid forward path is broken and noise is injected during particular intervals.

A number of researchers have proposed and evaluated continuously and noncontinuously adapting feedback-reduction algorithms (Bustamante *et al.*, 1989; Kates, 1991; Engebretsen and French-St. George, 1993; French-St. George *et al.*, 1993; Joson *et al.*, 1993; Smriga, 1993; Edwards, 1998; Kaelin *et al.*, 1998; Chi *et al.*, 1999; Kates, 1999). In general, they have obtained a 10–20-dB increase in the maximum stable gain for simulations or measurements on a mannequin. However, it is difficult to make comparisons across algorithms because of the different feedback paths used. Furthermore, studies including hearing-impaired listeners and realistic acoustic conditions either have been informal (Bustamante *et al.*, 1989) or have not assessed the algorithms' ability to increase the maximum stable gain (French-St. George *et al.*, 1993). In addition, little attention has been given to formal assessment of the algorithms' effects on sound quality. A notable exception is field trials of the DFS system demonstrating an 11-dB average increase in stable gain, with 15 of 21 hearing-impaired subjects preferring the experimental aid to their own (Smriga, 1993), and 5–10-dB increase in stable gain at 2 kHz, with 9 of 10 profoundly impaired subjects preferring the experimental aid to their own (Henningsen *et al.*, 1994).

In order to make valid comparisons among algorithms, Maxwell and Zurek (1995) evaluated five algorithms based largely on the feedback-reduction algorithms discussed above. They used three simulated feedback paths (two static and one dynamic) and two experienced, normal-hearing listeners. All of the algorithms studied provided added stable gains of 4–14 dB with mild detrimental effects on sound quality. Their results did not show any one algorithm to be

[a]Present address: Sensimetrics Corporation, 48 Grove Street, Somerville, MA 02144.
[b]Present address: CORE Physician Services, Exeter, NH.

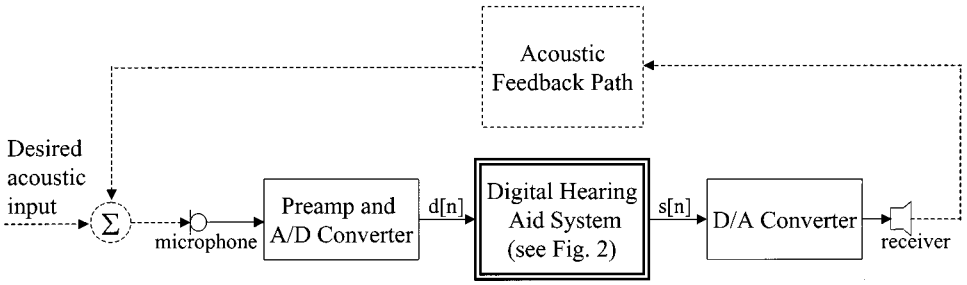


FIG. 1. Block diagram of a digital hearing aid system with acoustic feedback.

clearly superior, but did indicate that several feedback-reduction algorithms warrant further evaluation under more realistic conditions. The current work evaluates three promising feedback-reduction algorithms identified by Maxwell and Zurek.[1] An experimental digital hearing aid system with microphones and receivers worn by the subject permitted testing of the algorithms with *in situ* feedback paths. Measurements of stable gain and sound quality ratings by hearing-impaired listeners were collected for multiple vent configurations with actual dynamic feedback paths.

II. HEARING AID ALGORITHMS

This study considered three experimental feedback-reduction algorithms:

(1) Continuous, No Noise (CNN), similar to the continuous adaptive feedback oscillation (C-AFC) algorithm evaluated by Maxwell and Zurek (1995), based on the algorithm proposed by Bustamante *et al.* (1989);
(2) Open loop, Noise when Oscillation detected (ONO), similar to the noncontinuous adaptive feedback cancellation, oscillation triggered (NC-AFC-OT) algorithm evaluated by Maxwell and Zurek, based on the algorithm proposed by Kates (1991);
(3) Open loop, Noise when Quiet or oscillation detected (ONQ), similar to the noncontinuous adaptive feedback cancellation, quiet interval (NC-AFC-QI) algorithm proposed by Maxwell and Zurek (1995).

All of the feedback-reduction algorithms considered in this study operate in conjunction with a hearing aid forward path that processes the microphone signal for presentation to the receiver. In the absence of any feedback-reduction algorithms, this hearing aid forward path comprised a reference algorithm (REF) for comparison with the three experimental feedback-reduction algorithms. The hearing aid forward path is described in detail in Sec. II D.

A. The CNN algorithm

Figure 2(a) shows a block diagram of the CNN algorithm. This algorithm attempts to estimate the acoustic feedback path by continuously adapting to minimize the difference between the hearing aid input, $d[n]$, and a digitally filtered version of the hearing aid output, $s[n]$. The delayed hearing aid output, $x[n]$, passes through an adaptive filter. The adaptive filter output is subtracted from the delayed hearing aid input, producing $e[n]$. In the ideal case, the combination of the delay, Δ_x, and the steady-state adaptive filter weights, $\mathbf{w}[\infty]$, will exactly match the overall feedback path from $s[n]$ to $d[n-\Delta_{IN}]$, which includes the acoustic feedback path plus the effects of the preamplifier, the A/D and D/A converters, and the input delay, Δ_{IN}. In practice, the

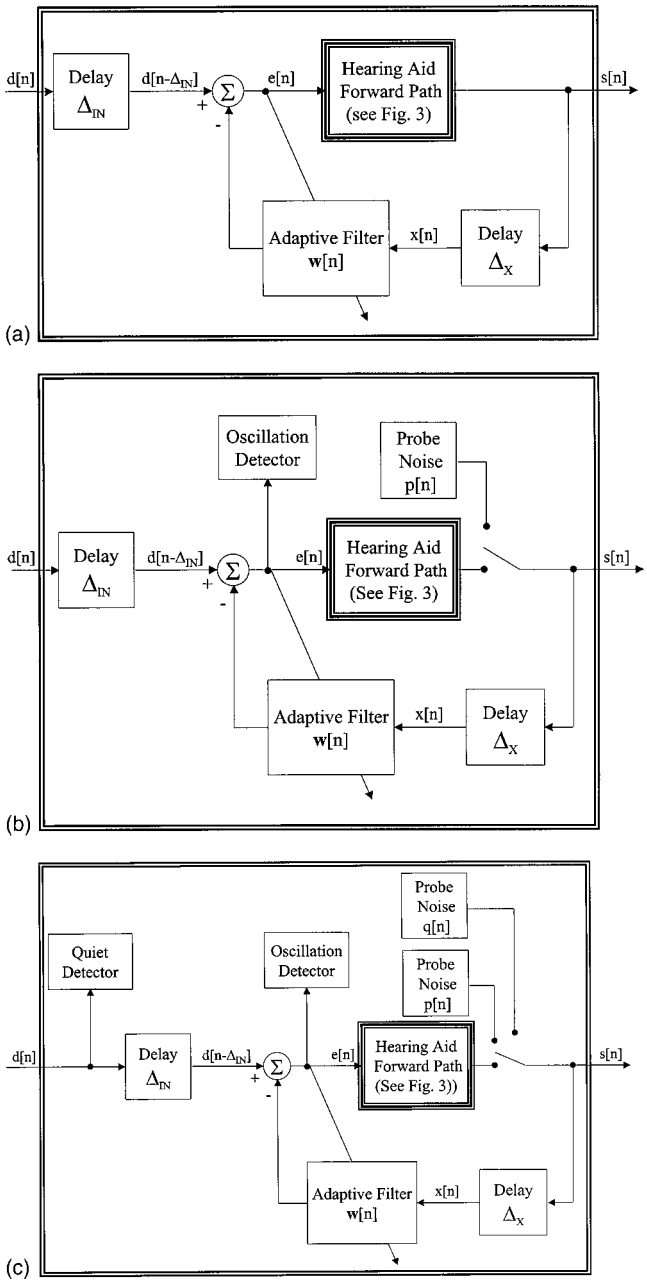


FIG. 2. Block diagrams of three experimental feedback-reduction algorithms described in the text. (a) CNN algorithm, (b) ONO algorithm, (c) ONQ algorithm.

adaptive filter weights only estimate the overall feedback path, and $e[n]$ is the error signal with respect to this estimation. However, with respect to the operation of the hearing aid, $e[n]$ is the desired signal plus residual error. In the absence of acoustic feedback, $d[n]$ (and therefore $e[n]$) are digitized versions of the desired acoustic input, that is, the ambient signal at the microphone.

The adaptive filter weights are updated using the least-mean-square (LMS) algorithm (Widrow and Stearns, 1985) using the equation

$$\mathbf{w}[n+1]=\mathbf{w}[n]+\mu[n]e[n]\mathbf{x}[n], \tag{1}$$

where boldface characters represent vectors, $\mathbf{w}[n]$ is the $L\times 1$ vector of adaptive filter weights, L is the length of the adaptive filter, $e[n]$ is the error signal described above, and $\mathbf{x}[n]$ is the data vector containing L samples of the hearing aid output present in the adaptive filter's tapped delay line at time n. The time-varying adaptive step size $\mu[n]$ is computed using the sum method (Greenberg, 1998), given by

$$\mu[n]=\frac{\alpha}{L(\sigma_e^2[n]+\sigma_x^2[n])}, \tag{2}$$

where α is the dimensionless step-size constant and $\sigma_e^2[n]$ and $\sigma_x^2[n]$ are power estimates computed from $e[n]$ and $x[n]$, respectively, by squaring and then smoothing with a first-order infinite-impulse response (IIR) low-pass filter with time constant equal to L. This method of normalizing the adaptive step size based on the sum of the powers in $e[n]$ and $x[n]$ is appropriate for this application because $e[n]$ is the residual error plus the desired signal, which is relatively strong and presumably uncorrelated with the feedback signal being estimated by the adaptive filter.

The CNN algorithm evaluated in this study used the following parameter values: $\Delta_{IN}=6$ ms, $\Delta_x=10$ ms, $L=160$, and $\alpha=0.005$.

1. Stability of the CNN algorithm

Because of the recursive configuration of the adaptive filter in the CNN algorithm, the adaptive filter weights may fail to converge to a global minimum and may become unstable (Maxwell and Zurek, 1995). Two specific situations have been observed to lead to unstable behavior. In both situations, a strong correlation exists between the error signal, $e[n]$, and the adaptive filter data vector, $\mathbf{x}[n]$, and the adaptive filter weights are not able to reduce this correlation. In most cases, instability can be avoided by appropriate selection of the delays Δ_x and Δ_{IN}, as discussed below.

In the first situation, the correlation is inherent in the ambient signal itself (as in the case of a sinusoidal system input) and cannot be reduced because of the short length of the adaptive filter. If the adaptive filter length is sufficiently long, cancellation of the ambient signal results instead of instability. Overcoming this problem requires insertion of a delay (Δ_x) between the hearing aid output, $s[n]$, and the adaptive filter input, $x[n]$, to decorrelate the desired signal component of $e[n]$ from the adaptive filter input (Bustamante *et al.*, 1989; Siqueira and Alwan, 2000). For this purpose, a delay of $\Delta_x=1$ ms has been shown to be sufficient (Bustamante *et al.*, 1989). However, steady-state tonal signals remain correlated even with such a delay, requiring constraints on the magnitude of the adaptive filter weights (Kates, 1999).

In the second situation, the adaptive filter weights are not able to reduce the correlation introduced by the acoustic feedback because improper selection of overall delays prevents the adaptive filter weights from approximating salient features in the feedback path. (Overall matching of the delays is also required to obtain good performance of the ONO and ONQ algorithms. However, mismatched delays do not cause instability, because the adaptive filters are not recursive in these algorithms with open-loop configurations.) For the algorithm shown in Fig. 2(a) combined with the system described by Fig. 1, the delays in the modeled feedback path (Δ_x plus $\mathbf{w}[n]$) should match the delays in the actual feedback path between $s[n]$ and $d[n-\Delta_{IN}]$. While the delays due to the acoustic feedback path are expected to be less than 1 ms (Bustamante *et al.*, 1989; Hellgren *et al.*, 1999), the system used in this study included sigma-delta oversampling A/D and A/D converters which added delays of 2.2 and 2.1 ms, respectively, when sampling at 16 kHz. As a result, the delay in the feedback path from $s[n]$ to $d[n-\Delta_{IN}]$ is equal to the delay introduced by the acoustic feedback path plus $\Delta_{IN}+4.3$ ms. Therefore, $\Delta_x=\Delta_{IN}+4$ ms is a reasonable choice. Although this study used values of $\Delta_{IN}=6$ ms and $\Delta_x=10$ ms, future work should assess the possibility of reducing the overall delay by testing the performance of the system with $\Delta_x=4$ ms and $\Delta_{IN}=0$.

Siqueira and Alwan (2000) suggest that placing the delay in the hearing aid forward path is preferable to the current placement in the cancellation path. This preference is based on their observation that the recursive nature of the continuously adapting algorithm leads to a bias in the Wiener solution of the adaptive filter weights; this bias is independent of the feedback path for delays placed in the forward path, while the minimum bias is limited by the first Δ_x samples of the actual feedback path impulse response for delays placed in the cancellation path. However, in the current implementation, the first Δ_x samples of the overall feedback path (as measured between $s[n]$ and $d[n-\Delta_{IN}]$) are guaranteed to be zero due to the delays introduced by the A/D and D/A converters and Δ_{IN}. Therefore, if the delays are matched as discussed above, placing the delay in the cancellation path does not contribute to bias in the adaptive filter weights.

B. The ONO algorithm

Figure 2(b) shows a block diagram of the ONO algorithm. This algorithm employs a detector to determine when strong oscillation occurs. When this condition is met, the normal hearing aid forward path is broken and a broadband probe noise, $p[n]$, is presented to the hearing aid receiver for a predetermined interval. The adaptive filter that estimates the feedback path is only updated during the interval when the probe noise is active. In theory, this should give better performance since it allows estimation of the acoustic feedback path without addition of signal from the forward path.

The disadvantages of this approach are twofold: the disruption of the hearing aid forward path and the potential disturbance of the probe noise itself.

An adaptive notch filter is used to detect oscillation. The adaptive notch filter is that specified by the first three lines of Kates' (1991) Eq. (1), with values of $\rho=0.7$ and $\mu=1.0$. The adaptive notch filter's input and output are each smoothed with a first-order IIR low-pass filter with 50-ms time constant. Then the detector computes the ratio of output power to input power. This is different than previous implementations (Kates, 1991; Maxwell and Zurek, 1995) which employed the difference between input power and output power. In this work the ratio was used to address the problems encountered with compression limiting described in Sec. II D. The ratio is then compared to a threshold of 0.05 (-13 dB); if the ratio is less than the threshold, the detector declares oscillation.

When oscillation is detected, probe noise is injected for a 2-s interval. The probe noise is zero-mean uniform random noise delivered at a level of 80 dB SPL in the ear canal. This interval is substantially longer than the 50 ms used previously (Kates, 1991; Maxwell and Zurek, 1995) but it was necessary to get good feedback-reduction performance in the current system. The level of the probe noise was fairly high and likely to be noticeable to most hearing-impaired listeners. Although the user may find it annoying, it is hoped that the noise will be less disruptive than the oscillation itself.

Although the adaptive filter weights are updated only during the 2-s probe noise, the details of the adaptation are the same as in Eqs. (1) and (2). The ONO algorithm evaluated in this study uses parameter values of $\Delta_x=4$ ms, $\Delta_{\mathrm{IN}}=0$ ms, $L=160$, and $\alpha=0.01$.

C. The ONQ algorithm

Figure 2(c) shows a block diagram of the ONQ algorithm. This algorithm is similar to the ONO algorithm, but in addition to the oscillation detector, it also employs a quiet-interval detector. The idea is that it may be possible to prevent oscillation by identifying quiet intervals due to brief pauses in the desired speech signal and then injecting a low-level probe noise, $q[n]$, and updating the adaptive filter weights during those intervals. If strong oscillation should occur despite the quiet-interval adaptation, then the algorithm will inject the higher-level probe noise, $p[n]$, and adapt for 2 s as in the ONO algorithm.

The quiet-interval detector is similar to the one described by Maxwell and Zurek. It operates by taking the absolute value of the samples at the hearing aid input and then filtering with a first-order IIR low-pass filter with a 25-ms time constant. This level is compared to a threshold selected to correspond to input levels of approximately 40 dB SPL. If all successive samples in a 10-ms interval are below this threshold, then the detector declares the start of a quiet interval. The quiet interval ends when a single sample exceeds a second threshold corresponding to approximately 45 dB SPL. Since the thresholds of the quiet detector are necessarily set to relatively low sound levels, this approach is only feasible in environments without substantial background noise. Adaptation of the filter weights, $\mathbf{w}[n]$, begins at the start of the quiet interval and continues for Δ_{IN} samples after the end of the quiet interval is detected. The quiet-interval detector operates on the hearing aid input signal, which is subsequently delayed by $\Delta_{\mathrm{IN}}=6$ ms, to permit the detector to use ''future'' information relative to the signal available to the feedback-reduction algorithm.

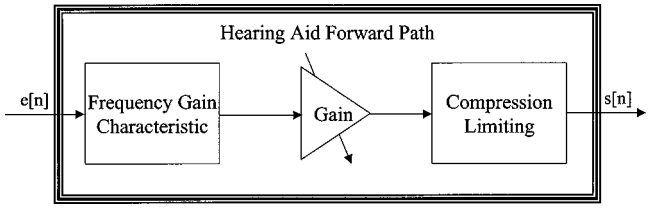


FIG. 3. Block diagram of hearing aid forward path common to all feedback-reduction algorithms.

The quiet-interval probe noise, $q[n]$, is turned on at the start of the quiet interval and turned off when the end of the quiet interval is detected. Subject-specific frequency-shaping was applied to $q[n]$ so that it paralleled each ear's absolute threshold. The level of this probe noise was adjusted so that octave-band levels matched absolute thresholds.

The ONQ algorithm uses the same method of oscillation detection and the same probe noise $p[n]$ as described in Sec. II B for the ONO algorithm. Although the adaptive filter weights are updated only during intervals when one of the two probe noises is active, details of the adaptation are the same as in Eqs. (1) and (2). The ONQ algorithm evaluated in this study uses parameter values of $\Delta_x=10$ ms, $L=160$, $\alpha=0.01$ when oscillation is detected, and $\alpha=0.005$ when quiet intervals are detected.

D. Common elements

The hearing aid forward path consists of the three components shown in Fig. 3: a frequency-gain characteristic, adjustable broadband gain, and compression limiting. The frequency-gain characteristic provides a frequency-dependent amplification that attempts to compensate for the subject's hearing loss. The method for fitting the frequency-gain characteristics is described in Sec. III C. The adjustable broadband gain, or volume control, is controlled by the experimenter. Section III D describes how the gain is adjusted to determine the maximum stable gain of each test condition. Compression limiting protects the subject from uncomfortably loud sounds by restricting the maximum level of the hearing aid output. The compression limiting used a 4-ms look ahead, 3-ms fast attack, and 100-ms slow release. Its input/output characteristic had unity slope below the compression threshold and infinite compression above the threshold. The compression limiting threshold was based on measurements of the subject's uncomfortable loudness levels (Cox, 1995).

Kaelin *et al.* (1998) discuss a number of issues related to interactions between adaptive feedback reduction and compression. In the current study, two problems were encountered with using compression limiting. First, the measurement of maximum stable gain is confounded by the compression limiter. If signal levels are high enough that compression is active, then increases in the broadband gain

are counteracted by attenuation in the compressor. This could lead to mistakenly large measurements of maximum stable gain. To prevent inflated measurements of this nature, the experimenter monitored a display showing the percentage of samples for which compression was active and then reduced the gain when the percentage exceeded 20%.

The second problem related to the performance of the oscillation detector used in the ONO and ONQ algorithms in the presence of compression limiting. The compression limiter can enter a cyclic mode initiated when the hearing aid begins to oscillate. This causes the compressor to lower the gain, thereby stopping the oscillation. When the oscillation stops, the compressor raises the gain and the oscillation resumes. This pattern resulted in a modulated oscillation with spectral characteristics that were difficult to detect with the adaptive notch filter as used previously. This problem was solved by widening the notch relative to previous work (Kates, 1991; Maxwell and Zurek, 1995) by setting the value of the parameter ρ to 0.7 and using the ratio of the notch filter input and output power, rather than the difference, as described in Sec. II B. This allowed the modulated oscillation to be detected, but the disadvantage of this solution was that the oscillation detector became slightly more prone to false alarms when no oscillation was present.

III. EXPERIMENTAL PROCEDURES

A. Hearing aid system

The hearing aid system used in this study consisted of microphones, preamplifier circuits, digital signal processing boards installed in a personal computer, and hearing aid receivers. For each ear, two Knowles EK-3024 omnidirectional microphones were mounted in a behind-the-ear (BTE) hearing aid shell with 12-mm spacing. The preamplified microphone signals were presented to the analog inputs of two DSP-96 boards (Ariel Corporation). Each of these boards contains 16-bit stereo A/D and D/A and a Motorola 96002 floating-point digital signal processing chip with 32-bit data paths and 40-MHz clock cycle. The boards were programmed to sample the omnidirectional microphone signals at 16 kHz, digitally combine the two microphone signals from each ear to simulate a simple directional microphone, and implement the feedback-reduction algorithms described in Sec. II. The analog output signal was presented to the subject via a Knowles receiver (either ED-1932 or CI-2748, depending on the subject's hearing loss) embedded in a custom full shell in-the-ear (ITE) module with a 1.5-mm diameter vent.

B. Test environment

The experiments were performed in a double-wall soundproof room with internal dimensions of $2.7\times2.5\times2.0$ m. The walls and ceiling were perforated metal panels with acoustically absorptive foam backing. The floor was a solid metal panel covered with a carpet. The space is relatively nonreverberant, but not anechoic. The subject, wearing the microphones and receivers, sat in the center of the booth. Cables connected the microphones and receivers to a preamplifier. A Radio Shack Optimus Pro 7 loudspeaker was located straight ahead of the subject at a distance of 1 m and a height of 1.1 m. The sound stimulus consisted of continuous speech material read by a male talker (Packard, 1995). The speech was presented at 60 dB SPL measured by a sound level meter (C-weighted, slow response) at the position of the subject's head with the subject absent from the booth. During the fitting procedure described in Sec. III C, the speech level was 65 dB SPL. The experimenter, the computer containing the DSP boards, and the test equipment were located outside the booth. The subject used a hand-held terminal to communicate with the experimenter and to record responses.

C. Hearing aid fitting

The experimental hearing aid system was programmed to implement two 127-point linear-phase finite-impulse response (FIR) filters to provide an appropriate frequency-gain characteristic for each ear. The goal of the fitting process is to determine binaural frequency-gain characteristic based on a compromise between matching the insertion gains prescribed by NAL-R (Byrne and Dillon, 1986) and matching the insertion gains provided by the subject's personal hearing aids. The frequency-gain characteristics are based on

(a) The NAL-R target insertion gains, computed from measurements of the subject's unmasked thresholds for each ear,
(b) real-ear insertion gains of the subject's personal hearing aids with the volume control set by the subject for comfortable listening of conversational speech in quiet,
(c) real-ear insertion gains of the experimental hearing aid system with known flat digital gain for frequency response,
(d) feedback path transfer functions computed by breaking the hearing aid forward path, injecting a known random noise sequence in place of $s[n]$, and measuring the received digital signal $d[n]$ in the absence of any desired acoustic input.

These quantities are all functions of frequency and are used to create two initial binaural frequency-gain characteristics. The first one matches the insertion gains of the subject's personal hearing aids (PHA). The second one matches either PHA or the insertion gain prescribed by NAL-R, whichever is greater at each frequency. The rationale for this choice was that subjects were likely to accept frequency-gain characteristics that matched their own aids, and might accept additional gain beyond that provided by their personal aids when prescribed by NAL-R. It was assumed that subjects would not accept reductions in gain relative to their personal aids, even if those reductions were prescribed by NAL-R. All of the frequency-gain characteristics are subject to the constraint that the gain at each frequency is limited to 5 dB below the level resulting in instability as predicted by the magnitude of the feedback path transfer functions.

During the fitting procedure, the subject listens binaurally to conversational speech, first via the PHA responses in order to verify that it ''sounds like'' their own aids. Next, the subject listens via the maximum PHA/NAL responses and is

asked if that would be acceptable for listening for extended periods. If the subject accepts, then the maximum PHA/NAL response is used for the study. If not, a new frequency-gain characteristic halfway between PHA and maximum PHA/NAL is created and the subject is asked again. This process continues until the subject accepts a set of binaural frequency-gain characteristics, or until the difference between successive frequency-gain characteristics is less than 5 dB for both ears and all frequencies. The result of this process is that the subject receives the strongest acceptable binaural frequency-gain characteristics with maximum gains at each frequency limited by the insertion gains of either their personal hearing aids or the NAL-R target.

During the fitting procedure, the vent in the ITE module was reduced to 0.5 mm with a vent plug when possible. However, for some subjects, the desired gains could only be provided without feedback when a closed plug was used to completely occlude the vent.

D. Determination of maximum stable gain

For this study, the maximum stable gain (MSG) of a hearing aid algorithm is defined as the largest value such that

(a) the hearing aid does not oscillate,
(b) algorithm artifacts (such as distortion or repeated injection of probe noise) do not substantially degrade the target speech,
(c) compression limiting occurs on fewer than 20% of samples,[2] and
(d) a 30-dB gain limit is not exceeded.

The hearing aid output signal delivered to the receiver in the subject's ITE module was also available for the experimenter to monitor via headphones. A gain of 0 dB corresponds to the frequency-gain characteristic determined as described in Sec. III C, and nonzero gain values refer to broadband amplification or attenuation of the fitted frequency-gain characteristic provided by the adjustable gain element depicted in Fig. 3. The experimenter used an interface that allowed control of the gain in 1-dB increments and displayed the percent of samples on which compression limiting was applied. The gain value was initially set to −10 dB, and slowly increased until reaching a limit imposed by one of the four conditions listed above.[3] If necessary, the gain was reduced slightly to stop oscillation or to reduce artifacts to a level judged to be tolerable by the experimenter. The subject was then instructed to move his/her jaw. If jaw motion caused the hearing aid to oscillate, the gain was reduced and further jaw motion requested. This procedure continued until the experimenter determined the MSG, the gain at which jaw motion did not trigger oscillation and the other three conditions were not violated.

E. Ratings

After the MSG was determined, the subject was presented with speech through the hearing aid system with the gain set to four levels: 0 dB, MSG, MSG−3 dB, and MSG−6 dB.[4] At each gain level, the speech stimulus was presented for 20 s, and the subject was instructed to move his/her jaw during that time. After each stimulus presentation, the subject was asked to provide subjective ratings of pleasantness and intelligibility using the scales shown in Fig. 4.

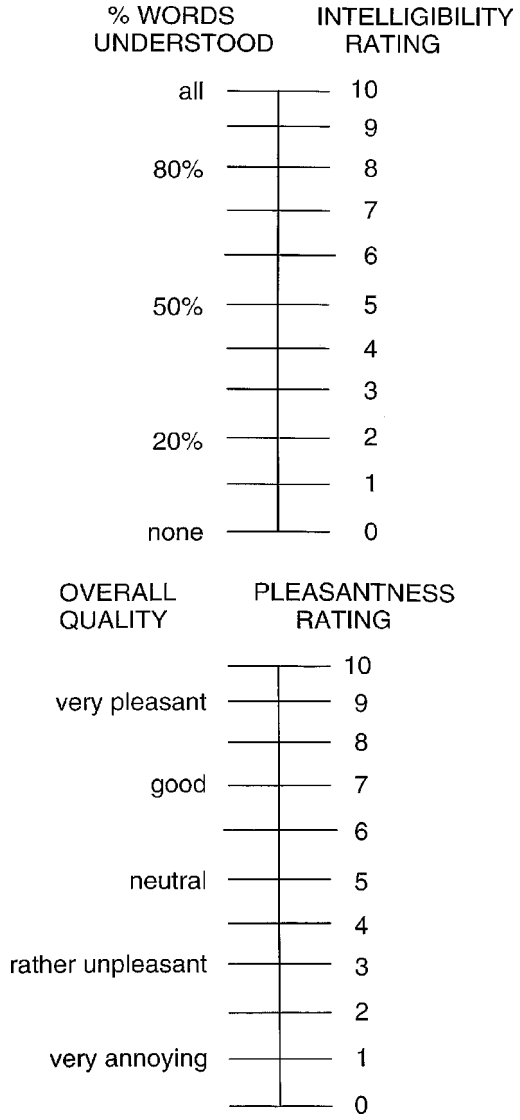


FIG. 4. Scales used by subjects for rating pleasantness and intelligibility.

F. Subjects

Table I describes the seven subjects who participated in this study, and Fig. 5 shows their audiograms. Two additional subjects could not participate because the compression limiting threshold dictated by their UCLs was sufficiently low that the effective gain of the aids could not be increased

TABLE I. Description of subjects.

Subject	Gender	Age	Receiver
JH	F	45	CI
HK	F	71	CI
GI	M	63	ED
PG	M	58	CI
TM	M	72	ED
HB	M	71	ED
WF	M	77	ED

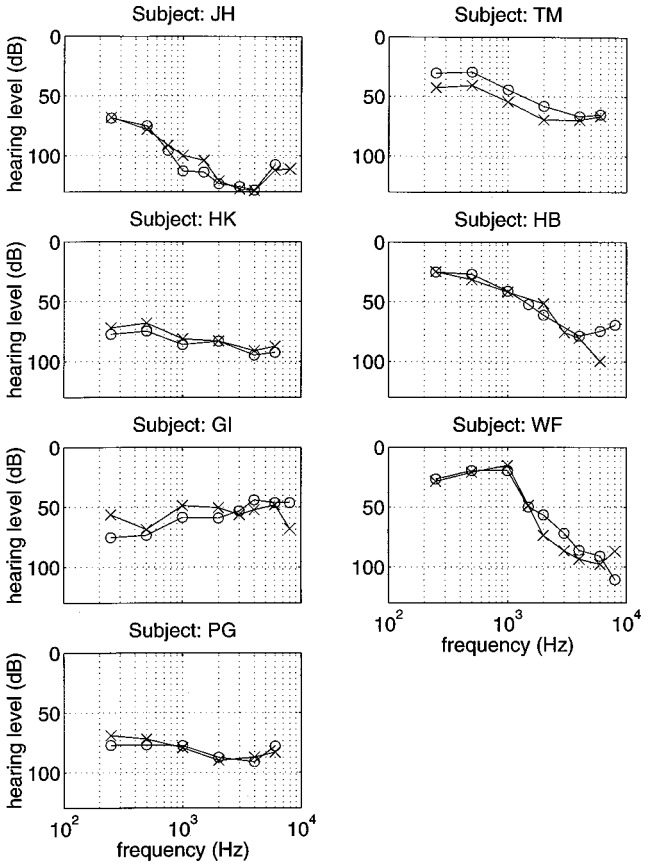


FIG. 5. Audiograms of the seven subjects participating in this study.

enough to cause oscillation. All subjects had bilaterally symmetric sensorineural hearing loss and passed tympanometry screening, except for one subject (GI) whose right ear had a mild conductive component below 1000 Hz and showed reduced compliance. All subjects had at least four years experience wearing hearing aids binaurally.

G. Test conditions

Each ear was tested individually with the untested ear occluded by the ITE module with no signal input to the receiver. Two vent conditions were tested for each ear: in the *open* vent condition, the 1.5-mm vent was left open, while in the *closed* vent condition a vent plug was used to either reduce the vent to 0.5 mm (subjects TM, HB, WF, GI) or to completely occlude the vent (subjects JH, HK, PG). This gave a total of four ear/vent combinations (right/closed right/open left/closed left/open) for each subject. Each trial tested a single ear/vent combination and included eight runs comprised of two repetitions[5] of each of the four algorithms (CNN, ONO, ONQ, and REF) described in Sec. II. The order of the runs was randomized within each trial. A single run consisted of determining the MSG as described in Sec. III D and collecting the eight ratings as in Sec. III E. The order of presentation of the gains used for rating was randomized within each run. A typical trial determining eight MSGs and collecting two ratings for each of 32 presentations lasted 40–50 min. One or two trials were performed during a subject visit, with subjects returning for up to four visits to complete testing of the four ear/vent combinations.

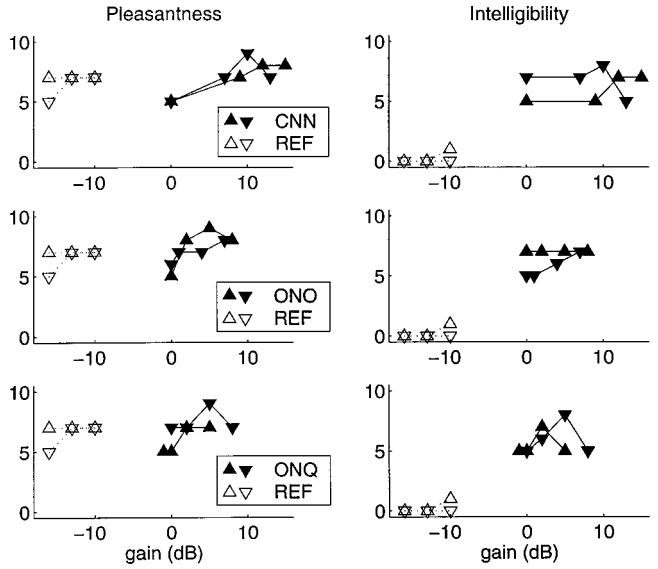


FIG. 6. Ratings of pleasantness and intelligibility vs gain for subject HK's left ear in the open vent condition. Each panel compares one experimental algorithm to the reference algorithm. Each series of connected points represents a single run.

IV. RESULTS

Before presenting the subject-averaged data, two complete sets of raw data for different subjects and ear/vent combinations will be described. These examples will illustrate trends in the ratings as the gain is varied and provide insight into the average results. These particular data were selected because they illustrate two extremes in terms of the algorithms' effects on subjective ratings, related to the subject's need (or lack thereof) for additional gain in order to make the speech intelligible.

Figures 6 and 7 each show results for one subject and one ear/vent combination. Each series of connected points shows the ratings as a function of gain for a single run. The MSG determined for a run is represented by the gain value associated with the rightmost data point in a connected series. Any benefit provided by an experimental algorithm

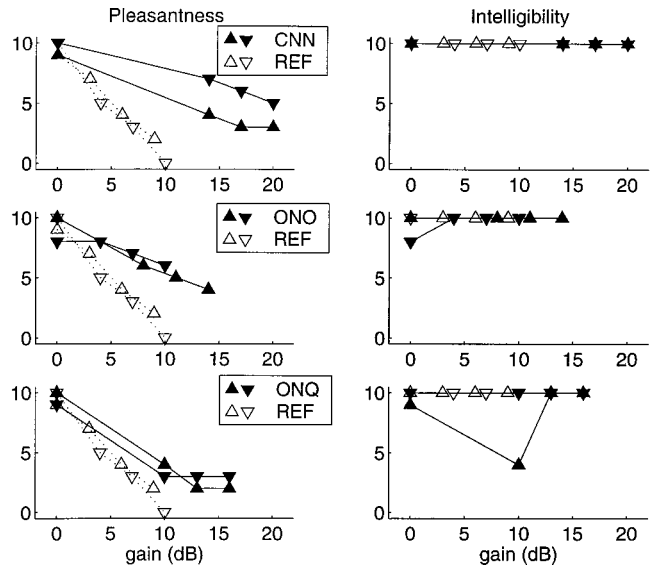


FIG. 7. Ratings of pleasantness and intelligibility versus gain for subject PG's left ear in the open vent condition, as in Fig. 6.

TABLE II. Added stable gain (ASG) averaged over repetitions for each experimental algorithm, subject, and ear/vent condition.

Subject	Algorithm	Closed/left	Closed/right	Open/left	Open/right
JH	CNN	5.5	10.5	25.0	15.5
	ONO	4.5	8.5	17.0	11.0
	ONQ	4.0	9.5	17.5	12.0
HK	CNN	10.5	15.0	24.0	18.0
	ONO	3.0	2.5	17.5	11.0
	ONQ	0.5	9.5	16.5	12.5
GI	CNN	6.0		17.5	11.5
	ONO	3.0		9.5	0.0
	ONQ	3.0		5.5	0.0
TM	CNN	3.5	2.0	2.0	5.5
	ONO	2.0	1.0	−0.5	6.0
	ONQ	1.5	2.0	0.5	3.0
PG	CNN	1.5	9.8	10.5	13.5
	ONO	0.0	6.8	2.5	10.0
	ONQ	2.5	6.8	6.5	11.0
HB	CNN	2.0	−1.0	5.5	2.0
	ONO	0.0	3.0	0.5	−1.0
	ONQ	−1.0	0.5	1.5	−0.5
WF	CNN	0.0	2.0	3.0	9.3
	ONO	0.0	1.5	1.5	9.7
	ONQ	0.0	1.0	1.5	7.0

(filled triangles) over the reference (open triangles) is indicated either by a greater MSG or by a higher rating at the same gain. The amount by which MSG for an experimental algorithm exceeds MSG for the reference algorithm in the same ear/vent condition is the added stable gain (ASG).

Figure 6 shows results for subject HK's left ear in the open vent condition. Averaging over repetitions, the CNN, ONO, and ONQ algorithms provided ASGs of 24.0, 17.5, and 16.5 dB, respectively. For the reference algorithm, the MSG was −10 dB, well below the subject's preferred gain setting of 0 dB. For this subject, the desired gain could only be provided in the closed vent condition, when a closed plug was used to completely occlude the vent. For the open vent condition, the reference algorithm was unstable at the preferred gain setting, resulting in a negative MSG value. The poor intelligibility ratings for the reference algorithm reflect the fact that the desired gain could not be provided. In this case, the benefit of the feedback-reduction algorithms is obvious; by permitting higher gains in the open vent condition, they increase audibility and intelligibility with a more comfortable earmold fitting. The companion pleasantness ratings indicate no substantial difference in pleasantness between the experimental algorithms and the reference algorithm.

Figure 7 shows results for subject PG's left ear with an open vent. Averaging over repetitions, the CNN, ONO, and ONQ algorithms provided added stable gains of 10.5, 2.5, and 6.5 dB, respectively. Although the subject judged almost all presentations to be highly intelligible, he found that the speech stimuli became unpleasant as the gain increased, presumably due to ''ringing'' as the hearing aid approached its maximum stable gain. Comparing each feedback-reduction algorithm to the reference algorithm shows that the experimental algorithms were more pleasant at the same gain settings, or, equivalently, permitted higher gains to be used before becoming equally unpleasant.

The MSG values determined using the procedure described in Sec. III D were quite consistent across repetitions; the MSG values were within 1 dB on 60% of conditions and within 3 dB on 86% of conditions.

TABLE III. Maximum stable gain (MSG) for each algorithm, averaged over subjects, ear/vent combinations and repetitions. Underlined groupings are significantly different at $P=0.01$ by Tukey HSD test.

REF	ONO	ONQ	CNN
7.6 dB	12.5 dB	12.8 dB	16.1 dB

Table II summarizes the ASG values, averaged over repetitions, for each subject, algorithm, and ear/vent combination. Positive values indicate the increase in stable gain provided by the experimental algorithm relative to the reference algorithm. In most cases, the experimental algorithms provided an increase in the stable gain delivered by the hearing aid. In the few cases with negative ASG values, the reduction in stable gain was minimal (1 dB or less), and within the variability of measurements from the experimental procedure.

A repeated-measures analysis of variance (ANOVA) on the MSG data analyzed three fixed factors: subject, algorithm, and repetition number with ear/vent combination nested within subjects. Significant main effects of algorithm ($P<0.0001$) and subject ($P=0.0006$) were found, as well as a significant subject×algorithm interaction ($P=0.0001$).

Although the feedback-reduction algorithms provide more benefit for some subjects than for others, the overall trends show that the CNN algorithm consistently provides the highest MSG values, while the reference algorithm provides the lowest MSG values. This is confirmed by a *post hoc* Tukey HSD test of the MSG data, which showed significant differences ($P=0.01$) between CNN and all other algorithms and between REF and all other algorithms, but no significant difference between the ONO and ONQ algorithms. Table III shows the Tukey groupings of the mean MSG values, which indicate that, averaged over all other factors, the CNN algorithm provided the largest increase in stable gain relative to the reference algorithm (ASG

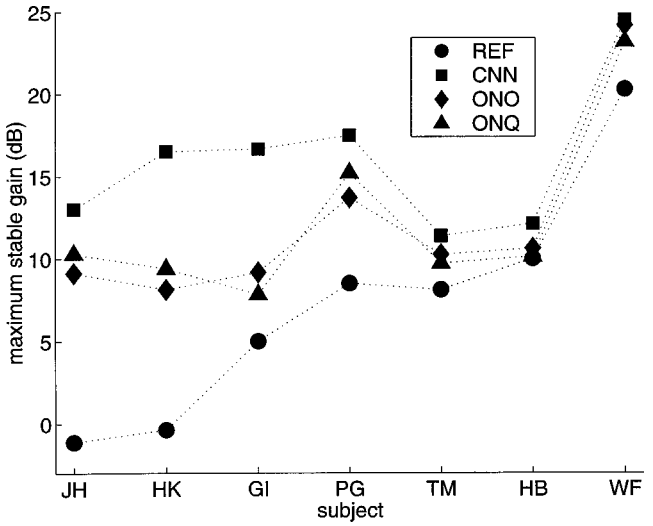


FIG. 8. Maximum stable gain (MSG) versus subject for each algorithm, averaged over ear/vent combinations and repetitions.

$=8.5$ dB) while the ONO and ONQ algorithms provided more moderate but significant increases in stable gain (ASG$\approx$5 dB).

To display the nature of the subject effect and the subject×algorithm interaction, Fig. 8 shows the MSG versus subject for each algorithm, averaged over ear/vent combinations and repetitions. The significant subject effect is expected and merely indicates that the amount of additional gain (relative to the subject's fitted frequency-gain characteristic) that can be applied before the hearing aid becomes unstable varies with subject. The significant subject×algorithm interaction indicates that not all subjects received the same benefit from the experimental algorithms. Figure 8 reveals that four subjects (JH, HK, GI, and PG) received fairly large benefits: an average ASG of 13 dB for the CNN algorithm and 7–8 dB for the ONO and ONQ algorithms. Three other subjects (TM, HB, WF) received an average ASG of 3 dB for the CNN algorithm and no significant benefit (based on repeating the ANOVA using only data from these three subjects) from the ONO and ONQ algorithms.

Because three of the four subjects receiving fairly large benefits (JH, HK, PG) used the CI receiver, the ANOVA was repeated with the receiver type as an additional factor. The ANOVA revealed no main effect of receiver, but significant receiver×algorithm interaction.

A repeated-measures multivariate analysis of variance (MANOVA) on the eight ratings collected in each run (intelligibility and pleasantness at four gain levels) found no significant algorithm effect using the same model described above for the ANOVA. This analysis compared ratings across algorithms at the same gain (preferred gain setting of 0 dB) as well as at various gains relative to MSG (that is, MSG, MSG-3 dB, and MSG-6 dB). In this case, the lack of an algorithm effect indicates that overall the subjects found all algorithms equally pleasant and equally intelligible: (1) at their preferred gain setting of 0 dB, and (2) at various gains relative to the maximum stable gain for that algorithm.

V. DISCUSSION

A. Stable gain

For the subjects and ear/vent combinations considered in this study, the CNN algorithm provided the largest increases in maximum stable gain relative to the reference algorithm. Averaged over all other factors, the CNN algorithm provided an average ASG of 8.5 dB, while the ONO and ONQ algorithms provided an average ASG of roughly 5 dB.

The benefit provided by the experimental algorithms was not uniform across subjects. The four subjects receiving the most benefit received an average ASG of 13 dB for the CNN algorithm and 7–8 dB for the ONO and ONQ algorithms. Three of those four subjects required relatively strong amplification, leading to use of the CI receiver, which has higher maximum output levels than the ED receiver used by the other subjects. The ANOVA including receiver type as a factor showed significant receiver×algorithm interaction. This result is difficult to interpret because the analysis is confounded by the fact that this receiver was used with subjects having more severe hearing losses. From the current data, it cannot be conclusively determined whether the significant interaction is the result of the receiver type itself or related to the degree of hearing loss. However, it is surmised that this effect is due to the hearing loss, not receiver type, because the fitting procedure designed frequency-gain characteristic filters to match specified target insertion gains, thereby compensating for the receivers' frequency responses. Therefore, subjects with more severe losses may receive greater benefits from the feedback-reduction algorithms. This is encouraging, because subjects with more severe losses require more amplification, making their hearing aids more prone to oscillation. However, a larger sample of subjects with varied audiograms would be required to conclusively predict the relationship between hearing loss and benefit from the experimental algorithms.

As shown in Table II, the experimental algorithms generally provided larger ASG values for the open vent conditions than for the closed vent conditions. This can be attributed to two factors. First, the initial fitting of the frequency-gain characteristic was performed in the closed vent condition, and the frequency-gain characteristic was modified to keep the gain at least 5 dB below the predicted level of instability at each frequency as determined by measurement of the feedback path. This resulted in a frequency-gain characteristic that was already limited to prevent oscillation, which in turn led to conservative measurements of the ASG for the closed vent conditions. Second, the larger vent increases feedback from the receiver to the microphone, giving the experimental algorithms more feedback to cancel, as it were, and leading to increased ASG values.

B. Subjective ratings

The MANOVA results indicate that relative to the reference algorithm, all three experimental algorithms provide increased stable gain with no significant effect on subjective intelligibility and pleasantness ratings. Several factors explain this lack of substantial algorithm effect on the intelligibility and pleasantness ratings.

Perhaps most surprising was the lack of algorithm effect for the ONO and ONQ algorithms, as it was expected that the injection of probe noise would lead to lower ratings for these algorithms. Preliminary listening by the experimenters confirmed the expectation that the probe noise is not only unpleasant, but that breaking the forward path to inject probe noise affects intelligibility. This concern was reinforced by informal comments from the subjects, who clearly expressed their objections to the probe noise. This dissatisfaction was not reflected in the subjective ratings because of the nature of the test procedure and the instructions given to the subjects. During most runs, the probe noise was injected several times as the gain was increased during the determination of MSG. During the ensuing 20-s evaluation intervals, the oscillation detector generally did not trigger, and probe noise was not injected, not even with the subject's jaw movements. Since subjects were instructed to provide ratings based solely on the evaluation interval, and the probe noise was rarely injected during that interval, the ratings do not reflect subject

displeasure with probe noise in the ONO and ONQ algorithms.

Additional factors explaining the lack of an algorithm effect relate to the nature of the values being compared. The experimental design included ratings collected at several gains relative to MSG for each algorithm. The lack of an algorithm effect for these ratings indicates that the subjects found the stimuli equally pleasant and intelligible *at levels relative to the maximum stable gain for that algorithm*. This result must be interpreted carefully, because the ratings are being compared across different gain levels. For example, in Fig. 7, the subject gave the reference algorithm pleasantness ratings of 3 at MSG−3=7 dB and 4 at MSG−3=6 dB for two repetitions. For the CNN algorithm, this is compared to pleasantness ratings of 3 and 6 at MSG−3=17 dB for two repetitions. Although it cannot be quantified from the data available, it is likely that pleasantness ratings for the CNN algorithm at gains of 6 or 7 dB would exceed the reference algorithm ratings of 3 and 4. Together with the significant algorithm effect on MSG, this interpretation leads to the speculation that subjects might have found the experimental algorithms *more pleasant at the same gain settings*, if they were substantially above 0 dB, as a corollary to the actual finding that the experimental algorithms were *equally pleasant while permitting higher gains to be used*.

The experimental design also included ratings collected at 0 dB in order to provide a point of comparison at the same absolute gain level. However, because of the precautions taken to reduce feedback in the initial fitting process, at 0 dB the hearing aid gain does not approach oscillation for most subjects and ear/vent conditions. When this is the case, the feedback reduction algorithms are inactive and not expected to affect pleasantness or intelligibility.

In the case of the intelligibility ratings, the lack of an algorithm effect at gain levels relative to the maximum stable gain for that algorithm is presumably because the subjects generally did not need additional gain beyond that provided by their fitted frequency-gain characteristic for the stimuli (60 dB SPL speech in quiet) presented in this study. (The results shown in Fig. 6 are a notable exception.) However, it is likely that the additional gain will improve intelligibility in other listening situations.

C. Relation to previous results

The ASG values determined in this study for actual dynamic feedback paths of hearing-impaired listeners are plausible in light of previous studies, which used mostly simulated and/or static feedback paths. The average ASG of 8.5 dB measured for the CNN algorithm is in accordance with the 6–10 dB reported by Bustamante *et al.* (1989) and the 7–14 dB reported by Maxwell and Zurek (1995).

Kates (1991) evaluated a version of the ONO algorithm using static feedback paths and reported ASGs of 17 dB for speech-shaped noise and 4–9 dB for a 1-kHz sine wave; Maxwell and Zurek (1995) used running speech and found ASGs of 12 dB for static feedback paths, but only 4 dB for a dynamic feedback path. In light of those results, the current finding of 5 dB of ASG for the ONO algorithm with dynamic feedback paths seems reasonable. The ONO implementation in this study used a longer adaptive filter and a smaller step-size constant ($L=160$ and $\alpha=0.005$) relative to the parameter values ($L=32,64$ and $\alpha=0.05,0.2$) used by Kates (1991). The current parameter values are expected to provide better steady-state performance at the expense of slower convergence relative to the previous implementations (Widrow and Stearns, 1985). These differences account for the longer probe noise interval (2 s instead of 50 ms) required in the current implementation. Using a shorter filter and larger step-size constant would permit a shorter probe noise interval, which might improve subjective sound quality judgements. However, such changes would not be expected to improve steady-state feedback cancellation performance or ASG values.

Maxwell and Zurek (1995) found comparable performance for the ONQ and ONO algorithms with static feedback paths (12 dB ASG), but an advantage for the ONQ algorithm with dynamic feedback paths (8 dB ASG) compared to the ONO algorithm (4 dB ASG). The current study found an average ASG of 5 dB for the ONQ algorithm, which was not significantly different from the ONO results. Any potential benefit of the ONQ algorithm relative to the ONO algorithm depends on accurate quiet interval detection. In Maxwell and Zurek (1995), the audio signal was delivered to the simulated hearing aid via an electrical connection, so the quiet interval detection operated on a clean signal in the absence of background noise. The current study used a quiet interval detector that was adjusted extensively for the background noise levels inside the soundproof room; even so, adaptation during injection of quiet interval probe noise did not improve performance. Therefore, despite Maxwell and Zurek's (1995) encouraging initial results with the ONQ algorithm, the current study did not reveal an advantage over the ONO algorithm, even in environments with relatively low levels of background noise.

The DFS system (Smriga, 1993) found 11 dB of ASG for dynamic feedback paths (Smriga, 1993), but all subjects in that study had severe to profound hearing losses. The current study found that subjects with the largest losses received greatest benefit from the algorithms; the average ASGs for the two subjects in this study (HK and JH) with losses comparable to the DFS subjects were 16, 9, and 11 dB for the CNN, ONO, and ONQ algorithms, respectively.

For all three algorithms, the adaptive filter length ($L=160$) was selected based on the limitations of the experimental hearing aid system. Althought this value is substantially longer than filters used in previous studies and what is presently practical in a commercial device, the relevant technology is evolving rapidly, and this study attempted to assess the potential benefits of feedback-reduction algorithms without restrictions based on the hardware presently available for implementing commercial devices. In theory, longer adaptive filters are expected to provide better steady-state feedback-reduction performance by creating a model of the feedback path that matches the actual feedback path more closely than a shorter filter. This closer match is due to two factors: the greater time span, which may improve performance in reverberation, and the higher frequency resolution provided by the longer filter, which may improve perfor-

mance in any acoustic environment. However, this speculation is not supported by the current results, which are comparable to previous results using much shorter filters [$L = 6,12$ in Bustamante *et al.* (1989), $L = 32,64$ in Kates (1991)]. Future work aimed at identifying factors limiting the performance of feedback-reduction algorithms should consider the effects of filter length and reverberation, as well as other likely factors such as internal noise and misadjustment of the adaptive filter weights.

D. Expected benefits

Based on the results of this study, the CNN algorithm is the logical choice for feedback reduction in hearing aids. The CNN algorithm provides the highest maximum stable gain values, and informal subject comments about the objectionable nature of probe noise support its choice over the ONO and ONQ algorithms. The CNN algorithm provides added stable gains up to 25 dB for some subjects, with no effect on speech quality.

The current evaluation of feedback-reduction algorithms was limited to speech signals in a relatively nonreverberant environment. The CNN algorithm is known to attenuate steady-state tonal signals (Bustamante *et al.*, 1989), but this problem can be solved by constraining the magnitude of the adaptive filter weights (Kates, 1999). Future work should assess the performance of the CNN algorithm with actual dynamic feedback paths in reverberation and in the presence of narrow-band signals.

While this study considered the effect of jaw movements, additional variations in feedback paths are expected when a telephone, hand, or other object is placed near the ear (Rafaely *et al.*, 2000). The CNN algorithm is expected to adapt to compensate for these greater fluctuations in the feedback path, but future work should quantify the algorithm's performance in such scenarios.

The CNN algorithm provides several benefits. First, current hearing-aid users will benefit from the ability to obtain the desired gain with more comfortable fittings due to looser earmolds and larger vents. Second, for individuals with severe and profound hearing losses, the CNN algorithm increases the possibility of obtaining successful fittings. Finally, it is expected that the CNN algorithm will improve intelligibility with no adverse effect on sound quality, by providing additional stable gain without ''ringing'' in many real-world listening situations.

ACKNOWLEDGMENTS

This work was supported by Contract No. N01-DC-5-2017 and by Grant No. NIH-2-R01-DC00117 from the National Institute of Deafness and Other Communicative Disorders. The authors are grateful to A. R. Brughera and W. M. Rabinowitz for their assistance in developing the experimental hearing aid system used in this study.

[1]Two other algorithms evaluated by Maxwell and Zurek were not considered in this study. The adaptive notch filter (ANF) algorithm was not included because it is only effective for narrow-band feedback paths. The continuous adaptive feedback cancellation with noise (C-AFC-N) algorithm, using continuous adaptation with probe noise, was not included because preliminary testing in this study revealed that it did not provide any improvements over a similar system without probe noise.

[2]This condition was required to prevent mistakenly large MSG measurements, as described in Sec. II D. The value of 20% was selected to detect moderate levels of compression on the peaks of the envelope of the speech signal, before the broadband gain was increased to a level where the problem described in Sec. II D was fully manifest.

[3]At any time, the subject could reduce the gain in 10-dB increments by pressing a button on the hand-held terminal. This was implemented as an additional precaution against uncomfortably loud sounds during hearing aid oscillation, even though the hearing aid output was already limited as described in Sec. II D.

[4]In conditions where the maximum stable gain was less than 0 dB, ratings were collected only at the latter three gain levels.

[5]In a few cases, extra runs were performed, resulting in three or four repetitions of all algorithms for a particular subject and ear/vent combination. For one subject, testing was suspended for one ear in the closed vent condition because the hearing aid system did not oscillate with the reference algorithm at 30 dB of gain.

Bustamante, D. K., Worrall, T. L., and Williamson, M. J. (**1989**). ''Measurement and adaptive suppression of acoustic feedback in hearing aids,'' Proc. Int. Conf. Acoustics, Speech, Signal Processing, ICASSP-89, pp. 2017–2020.

Byrne, D. J., and Dillon, H. A. (**1986**). ''The National Acoustic Laboratories (NAL) new procedure for selecting the gain and frequency response of a hearing aid,'' Ear Hear. **7**, 257–265.

Chi, H. F., Gao, S. X., and Soli, S. D. (**1999**). ''A novel approach of adaptive feedback cancellation for hearing aids,'' Int. Symp. on Circuits and Systems, ISCAS-99 III, pp. 195–198.

Cox, R. M. (**1995**). ''Using loudness data for hearing aid selection: The IHAFF approach,'' Hear. J. **48**(2), 10–44.

Edwards, B. W. (**1998**). ''Signal processing techniques for a DSP hearing aid,'' Int. Symp. on Circuits and Systems, ISCAS-98 VI, 586–589.

Engebretsen, A. M., and French-St. George, M. (**1993**). ''Properties of an adaptive feedback equalization algorithm,'' J. Rehabil. Res. Dev. **30**, 8–16.

French-St. George, M., Wood, D. J., and Engebretsen, A. M. (**1993**). ''Behavioral assessment of adaptive feedback equalization in a digital hearing aid,'' J. Rehabil. Res. Dev. **30**, 17–25.

Greenberg, J. E. (**1998**). ''Modified LMS algorithms for speech processing with an adaptive noise canceller,'' IEEE Trans. Speech Audio Process. **6**, 338–351.

Hellgren, J., Lunner, T., and Arlinger, S. (**1999**). ''System identification of feedback in hearing aids,'' J. Acoust. Soc. Am. **105**, 3481–3496.

Henningsen, L. B., Dyrlund, O., Bisgaard, N., and Brink, B. (**1994**). ''Digital feedback suppression (DFS). Clinical experiences when fitting a DFS hearing instrument on children,'' Scand. Audiol. **23**, 117–122.

Joson, H. A. L., Asano, F., Suzuki, Y., and Sone, T. (**1993**). ''Adaptive feedback cancellation with frequency compression for hearing aids,'' J. Acoust. Soc. Am. **94**, 3248–3254.

Kaelin, A., Lindgren, A., and Wyrsch, S. (**1998**). ''A digital frequency-domain implementation of a very high gain hearing aid with compensation for recruitment of loudness and acoustic echo cancellation,'' Signal Process. **64**, 71–85.

Kates, J. M. (**1991**). ''Feedback cancellation in hearing aids: Results from a computer simulation,'' IEEE Trans. Signal Process. **39**, 553–562.

Kates, J. M. (**1999**). ''Constrained adaptation for feedback cancellation in hearing aids,'' J. Acoust. Soc. Am. **106**, 1010–1019.

Maxwell, J. A., and Zurek, P. M. (**1995**). ''Reducing acoustic feedback in hearing aids,'' IEEE Trans. Speech Audio Process. **3**, 304–313.

Packard, D. (**1995**). The HP Way: How Bill Hewlett and I Built our Company (Harper Audio, New York, NY) (sound recording).

Rafaely, B., Roccasalva-Firenze, M., and Payne, E. (**2000**). ''Feedback path variability modeling for robust hearing aids,'' J. Acoust. Soc. Am. **107**, 2665–2673.

Siqueira, M. G., and Alwan, A. (**2000**). ''Steady-state analysis of continuous adaptation in acoustic feedback reduction systems for hearing aids,'' IEEE Trans. Speech Audio Process. **8**, 443–453.

Smriga, D. J. (**1993**). ''Digital signal processing to suppress feedback: Technology and test results,'' Hear. J. **46**(5), 28–33.

Widrow, B., and Stearns, S. D. (**1985**). *Adaptive Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ).

Speech recognition by normal-hearing and cochlear implant listeners as a function of intensity resolution

Philipos C. Loizou[a)]
Department of Electrical Engineering, University of Texas at Dallas, Richardson, Texas 75083-0688

Michael Dorman
Department of Speech and Hearing Science, Arizona State University, Tempe, Arizona 85287

Oguz Poroy
Department of Electrical Engineering, University of Texas at Dallas, Richardson, Texas 75083-0688

Tony Spahr
Department of Speech and Hearing Science, Arizona State University, Tempe, Arizona 85287

(Received 8 April 2000; accepted for publication 28 August 2000)

The importance of intensity resolution in terms of the number of intensity steps needed for speech recognition was assessed for normal-hearing and cochlear implant listeners. In experiment 1, the channel amplitudes extracted from a six-channel continuous interleaved sampling (CIS) processor were quantized into 2, 4, 8, 16, or 32 steps. Consonant recognition was assessed for five cochlear implant listeners, using the Med-El/CIS-link device, as a function of the number of steps in the electrical dynamic range. Results showed that eight steps within the dynamic range are sufficient for reaching asymptotic performance in consonant recognition. These results suggest that amplitude resolution is not a major factor in determining consonant identification. In experiment 2, the relationship between spectral resolution (number of channels) and intensity resolution (number of steps) in normal-hearing listeners was investigated. Speech was filtered through 4–20 frequency bands, synthesized as a linear combination of sine waves with amplitudes extracted from the envelopes of the bandpassed waveforms, and then quantized into 2–32 levels to produce stimuli with varying degrees of intensity resolution. Results showed that the number of steps needed to achieve asymptotic performance was a function of the number of channels and the speech material used. For vowels, asymptotic performance was obtained with four steps, while for consonants, eight steps were needed for most channel conditions, consistent with our findings in experiment 1. For sentences processed though 4 channels, 16 steps were needed to reach asymptotic performance, while for sentences processed through 16 channels, 4 steps were needed. The results with normal-hearing listeners on sentence recognition point to an inverse relationship between spectral resolution and intensity resolution. When spectral resolution is poor (i.e., a small number of channels is available) a relatively fine intensity resolution is needed to achieve high levels of understanding. Conversely, when the intensity resolution is poor, a high degree of spectral resolution is needed to achieve asymptotic performance. The results of this study, taken together with previous findings on the effect of reduced dynamic range, suggest that the performance of cochlear implant subjects is primarily limited by the small number (four to six) of channels received, and not by the small number of intensity steps or reduced dynamic range. © *2000 Acoustical Society of America.* [S0001-4966(00)04911-0]

PACS numbers: 43.71.Ky, 43.71.Es [CWT]

I. INTRODUCTION

Normal-hearing listeners can detect amplitude changes over a 100-dB acoustic dynamic range with a high resolution. The total number of discriminable intensity steps for normal acoustic hearing, as calculated from Weber fractions reported by Schroder *et al.* (1994), was found to be over 80 (Nelson *et al.*, 1996). The large dynamic range coupled with fine intensity resolution and fine spectral resolution allows normal-hearing listeners to maintain high speech intelligibility in background noise and in extreme signal conditions lacking traditional spectral cues.

In contrast, cochlear implant listeners have a small dynamic range, typically 5–30 dB, and receive only a small number (four to six) of channels of spectral information through their device, despite the large number of electrodes stimulated (Fishman *et al.*, 1997; Wilson, 1997). The intensity difference limens (DLs) in electrical stimulation are comparable, and sometimes better, than the DLs in acoustic stimulation, depending on the intensity level. Implant listeners can detect 1–2-dB changes in intensity near threshold and 0.25–0.5-dB changes at higher intensity levels (e.g., Shannon, 1983; Nelson *et al.*, 1996; Pfingst *et al.*, 1983; Hochmair-Desoyer *et al.*, 1981; Dillier *et al.*, 1983). Although some implant listeners (e.g., Hochmair-Desoyer *et al.*, 1981) can detect smaller intensity changes than

a)Electronic mail: loizou@utdallas.edu

normal-hearing listeners, the difference is not enough to compensate for the large difference in dynamic range. The number of discriminable steps within the dynamic range is considerably smaller than the number of steps in acoustic hearing, and also varies considerably among subjects. Nelson *et al.* (1996) found that in some subjects the 30-dB dynamic range of speech was coded by fewer than 8 steps while in other subjects it was coded by less than 45 steps.[1] The three factors cited above, spectral resolution (number of channels), dynamic range, and intensity resolution (number of discriminable intensity steps across the dynamic range), are all factors that can potentially affect the performance of cochlear implant subjects. The number of channels available clearly affects the transmission of spectral cues (Dorman *et al.*, 1997; Loizou *et al.*, 1999; Shannon *et al.*, 1995), while the dynamic range and the number of discriminable steps may affect the coding of temporal-envelope cues. However, the extent to which these factors, separately and/or by interactions, affect speech understanding is not yet well understood.

We have addressed the effect of reduced dynamic range on speech understanding in a previous study (Loizou *et al.*, 2000). In that study, we simulated the effect of reduced dynamic range using normal-hearing listeners as subjects to avoid the confounding factors (e.g., surviving ganglion cells, electrode insertion depth, etc.) associated with cochlear implants. Signals were processed in a manner similar to a six-channel continuous interleaved sampling (CIS) processor (Wilson *et al.*, 1991) and output as the sum of sine waves with frequencies centered in the middle of each analysis band. The amplitudes of the sine waves were compressed to fit within a 6-, 12-, 18-, and 24-dB dynamic range. Our results showed a significant effect of compression for all test materials, although the effect of the compression was different for the three test materials (vowels, consonants, and sentences). Vowel recognition was affected the most by the compression, while consonant recognition was affected the least by the compression. Sentence recognition was moderately affected. Similar findings were reported by Zeng and Galvin (1999) with four implant listeners using the Nucleus-22 SPEAK speech processor. Two conditions of reduced dynamic range were examined. In the first condition, they reduced the electrode dynamic range to 25% of the full range by artificially raising the threshold levels to 75% of the dynamic range. In the second condition, they created a binary representation of the acoustic amplitudes by further decreasing the comfortable levels to 76% of the dynamic range. Their results showed that vowel recognition was significantly affected (although marginally) by the dynamic range reduction, whereas consonant recognition was not significantly affected.

The above studies demonstrated a significant effect, albeit mild, of dynamic range reduction on speech recognition. It is possible, however, that the number of discriminable steps available in the patient's dynamic range may be a more important factor for speech perception than the dynamic range itself. That is, a patient with a large dynamic range might only have a few discriminable steps, and, similarly, a patient with a small dynamic range might have a large number of steps. The variability in the number of steps among the eight implant subjects in the Nelson *et al.* (1996) study led us to wonder whether the performance of the poorly performing subjects was limited by a small number of steps. To gain a purchase on the answer to this question, we created, for presentation to implant patients, VCV syllables that had been quantized into 2–512 intensity steps. In this manner we experimentally controlled the intensity resolution available to the patients. At issue in experiment 1 was how many quantized steps are necessary for patients to achieve an asymptotic level of consonant recognition. We chose consonant recognition as our measure since amplitude envelopes are significant cues to consonant identity.

To better understand the role of intensity resolution on speech recognition one also needs to take into account the number of channels available. There might be a trade-off relationship between spectral resolution (number of channels) and intensity resolution (number of quantization steps). This hypothesis is based on our view (Dorman *et al.*, 1997) that when speech is processed through a small number of channels, the relative differences in across-channel amplitudes must be used to code frequency information. In this view, if intensity resolution were to be distorted, then speech recognition ought to decline. On the other hand, when speech is processed through a large number of channels, fine intensity resolution might not be needed, since the frequency information can be coded by the channels which have energy. These questions are investigated in experiment 2 with normal-hearing listeners, where we assess speech intelligibility as a function of number of channels and as a function of number of intensity steps. Normal-hearing listeners were used because the channels and steps manipulations cannot be independently controlled with implant listeners due to the many confounding factors associated with electrical stimulation. The results of experiment 2 could also be used to benchmark the performance of cochlear implant listeners, i.e., to assess whether cochlear implant listeners extract the same amount of information as normal-hearing listeners under similar conditions of reduced intensity resolution. To produce speech with varying degrees of intensity resolution, we synthesized speech as a linear combination of sine waves and quantized the amplitudes of the sinewaves to 2–32 levels. The intelligibility of vowels, consonants, and sentences was assessed as a function of spectral resolution and as a function of intensity resolution.

II. EXPERIMENT 1: ELECTRIC HEARING

A. Method

1. Subjects

The subjects were five postlingually deafened adults who had used a CIS processor for periods ranging from 3 to 4 years. Each of the patients had used a four-channel, compressed-analog signal processor (Ineraid) for at least 4 years before being switched to a CIS processor. The patients ranged in age from 40 to 68 years and they were all native speakers of American English. Biographical data for each patient are presented in Table I. Four patients were fitted with a six-channel CIS processor and one patient (S2) was fitted with a five-channel CIS processor.

TABLE I. Biographical data of the five cochlear-implant users who participated in this study.

Subject	Gender	Age (years) at detection of hearing loss	Age at which hearing aid gave no benefit	Age fit with Ineraid	Age at testing	Etiology of hearing loss	Score on H.I.N.T sentences in quiet	Score on NU-6 words in quiet
S1	M	20	46	63	68	unknown	88	46
S2	F	10	46	47	55	unknown	44	20
S3	M	5	43	48	58	unknown	92	43
S4	F	7	31	33	40	unknown/ hereditary	100	80
S5	F	23	48	51	57	unknown	100	71

2. Speech material

The consonant test consisted of 20 /vCv/ consonants in three vowel environments, /i a u/, produced by a single female speaker, and was taken from the consonant database recorded at the House Ear Institute (Shannon *et al.*, 1999). The 20 consonants were /b p d t g k f v s z ʃ ð t ʃ d ʒ m n r l j w/.

3. Experimental setup

This experiment was performed on our laboratory cochlear implant processor, which was based on the design of the Geneva/RTI/MEEI wearable processor (Francois *et al.*, 1994). Several modifications were made to the Geneva design, the most important of which was the addition of five current sources. The block diagram of the laboratory processor is shown in Fig. 1. The input analog circuit consists of an audio multiplexer that selects the source of the input signal to the processor, several fixed-gain amplifiers, one variable-gain amplifier (adjusted externally by a sensitivity knob), an antialiasing filter, and a 16-bit A/D converter. The sampling rate of the A/D converter is controlled by the DSP chip, and for this study it was fixed at 22 kHz. The cutoff frequency of the antialiasing filter was set at 6.7 kHz. Once the signal is digitized, it is transmitted to the Motorola DSP56002 chip, where it is processed through the CIS strategy (see description in the following section). The CIS outputs are finally fed through a SSI port to the current sources built around a digital-to-analog converter. Biphasic pulses are generated, with amplitudes equal to the CIS envelope outputs, and sent to the electrodes for stimulation. The pulse width as well as the stimulation rate was controlled through software. More information about the hardware of the laboratory processor can be found in Poroy and Loizou (2000).

4. Signal processing and amplitude quantization

Signals were first processed through a high-pass preemphasis filter (1200-Hz cutoff), with a 3-dB/octave roll-off, and then bandpassed into six frequency bands using sixth-order Butterworth filters. The center frequencies of the six bandpass filters were 461, 756, 1237, 2025, 3316, and 5428 Hz. The envelopes of the filtered signals were extracted by full-wave rectification and low-pass filtering (second-order Butterworth) with a 400-Hz cutoff frequency. The six envelope amplitudes A_i $(i=1,2,...,6)$ were mapped to electrical amplitudes E_i using a power-law transformation:

$$E_i = cA_i^p + d,$$

where c and d are constants chosen so that the electrical amplitudes fall within the range of threshold and most-comfortable level, and p is the power exponent. The power exponent p was set equal to -0.0001 to obtain a compression function similar to the logarithmic function found in the Med-El/CIS link device.[2] The electrical amplitudes were then systematically quantized into Q steps $(Q=2,4,8,16,32,512)$ to create six different conditions with varying degrees of intensity resolution. The 512-step condition corresponded to the number of steps used in our default CIS implementation, and was labeled the ''unquantized''

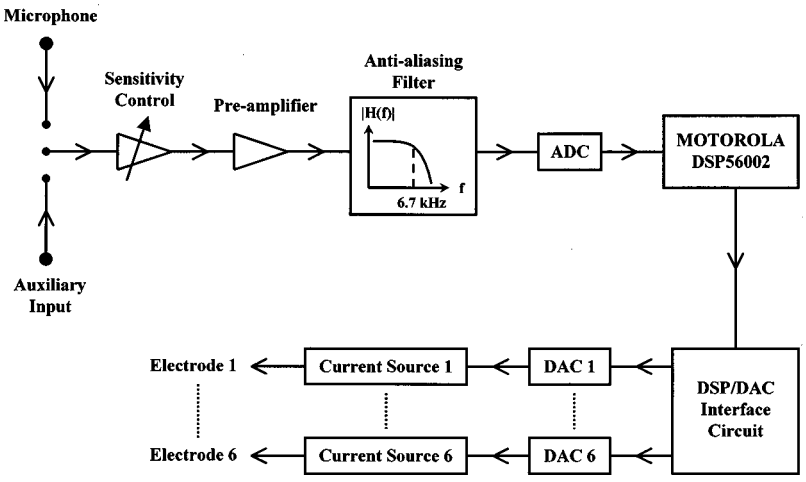


FIG. 1. Block diagram of the laboratory cochlear implant processor used in this study.

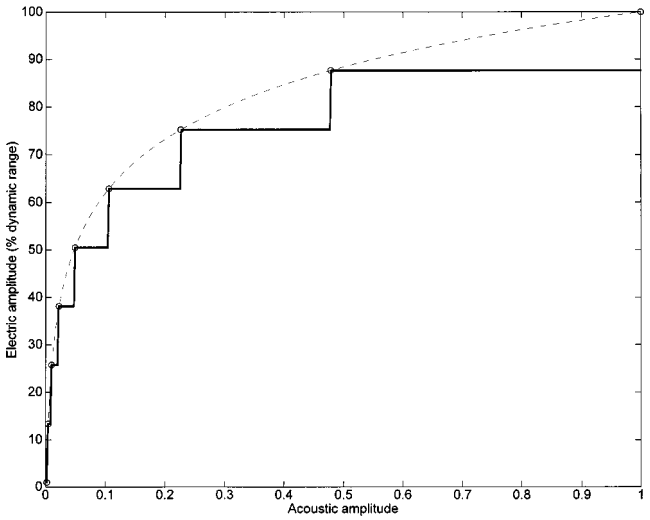


FIG. 2. Example of a logarithmic compression function quantized uniformly into eight steps. The solid curve shows the eight quantized levels and the dashed curve shows the original, unquantized, logarithmic function.

condition. [Note that there are some differences between the quantized steps of this study, and the psychophysical discriminable steps measured by Nelson *et al.* (1996) (see Discussion in Sec. II B]. As an example, Fig. 2 illustrates the quantization of the output dynamic range into eight uniform steps. The quantization step sizes, Δi, were estimated from the threshold (THR) and most-comfortable (MCL) levels of each electrode as follows:

$$\Delta_i = \frac{\mathrm{MCL}_i - \mathrm{THR}_i}{Q-1}, \quad i = 1,2,...,6,$$

where THR_i and MCL_i are the threshold and most-comfortable levels of the ith electrode, respectively. Note that since each electrode had different values for THR_i and MCL_i, the step sizes Δi were different for each electrode. The quantized envelope amplitudes were finally used to modulate biphasic pulses of duration 40 μs/phase at a stimulation rate of 2100 pulses/s. The electrodes were stimulated in the same order as in the subject's daily processors. For all but one subject, the electrodes were stimulated in ''staggered'' order.

5. Procedure

The test was divided into six sessions, one for each step condition. The six conditions were counterbalanced among subjects to avoid any order effects. There were nine repetitions of each consonant, presented in blocks of three repetitions each. The consonants were completely randomized. All test sessions were preceded by one practice session in which the identity of the consonants was indicated to the listeners.

The stimuli were presented to the subjects through a direct electrical connection using our laboratory processor at a comfortable listening level. To collect responses, a graphical interface was used that allowed the subjects to identify the words they heard by clicking on the corresponding button on the graphical interface.

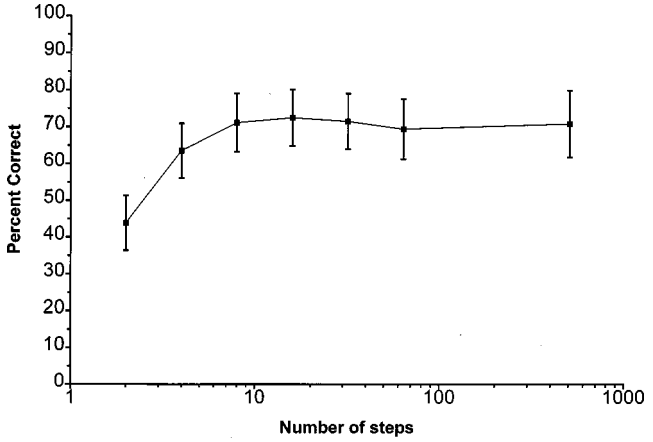


FIG. 3. Mean performance of cochlear-implant listeners on consonant recognition as a function of the number of quantized amplitude steps in the dynamic range. Error bars indicate ± standard errors of the mean.

B. Results and discussion

The results, scored in terms of percent correct, are shown in Fig. 3. Repeated measures analysis of variance showed a significant main effect [$F(6,24)=12.14$, $p<0.0005$] for the number of steps. *Post-hoc* analysis (Fisher's LSD—least significant difference) indicated that the asymptote in performance was obtained with eight steps, i.e., consonant recognition did not improve when eight or more steps were used.

The individual subjects' performances are shown in Fig. 4. The subjects' performances on consonant recognition varied considerably, from a low of 45% correct (subject S2) to a high of 90% correct (subject S4). For most subjects, performance improved substantially as the number of steps increased from two to eight, with only a small change in performance when more than eight steps were used. The fact that eight steps were sufficient for achieving asymptotic performance on the difficult test of consonant recognition demonstrates that a high degree of amplitude resolution is not necessary for consonant recognition. If eight steps are sufficient, then it is unlikely that intensity resolution limits many cochlear implant patients' performances on consonant recog-

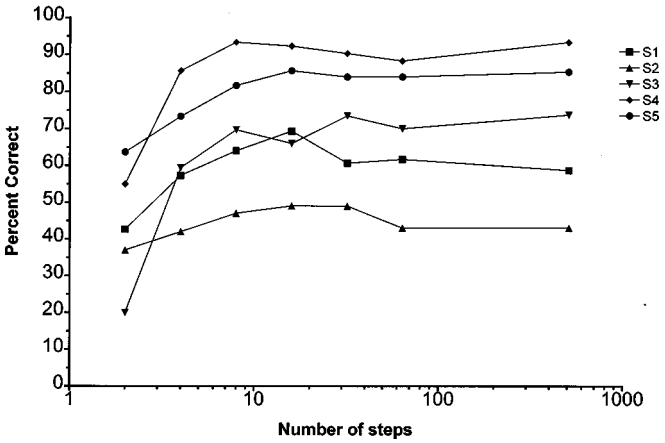


FIG. 4. Individual cochlear implant subjects' performance on consonant recognition as a function of the number of steps in the dynamic range.

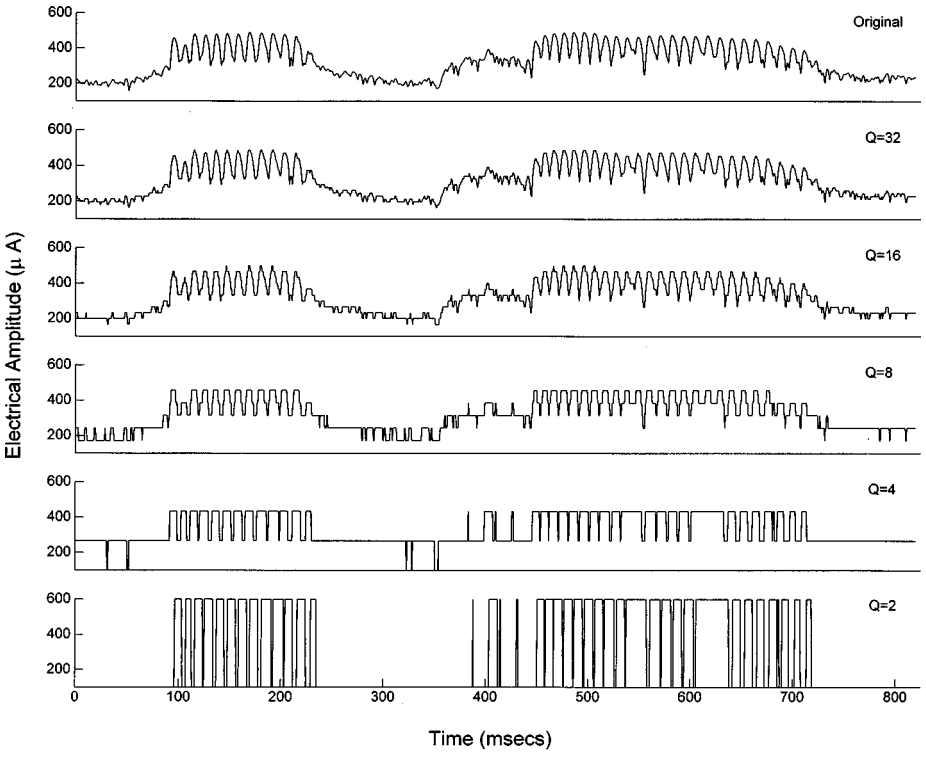


FIG. 5. Example of the waveform of the syllable /a p a/ quantized into 2, 4, 8, 16, and 32 steps. These waveforms were estimated by bandpass filtering speech into six frequency bands, computing the channel amplitudes in each band through envelope detection (400-Hz low-pass filter), and then quantizing the channel amplitudes into the indicated number of steps ($Q=2-32$). Only the (full-wave) rectified waveforms computed for channel 2 (center frequency=756 Hz) are shown in this figure.

nition since, judging from Nelson *et al.* (1996) and Zeng and Shannon (1999), most patients have at least eight discriminable steps in their dynamic range.

Wilson and colleagues (unpublished) have conducted an experiment similar in design to ours for a single Ineraid subject using a four-channel CIS processor. The subject was tested with a high rate (6900 pulses/s) and a low-rate (500 pulses/s) CIS processor. For the 24-consonant test (Tyler *et al.*, 1987), 8 steps were needed to attain asymptotic performance with the high-rate processor (consistent with our findings), and 16 steps were needed for the low-rate processor. Wilson's results suggest the possibility that the number of steps needed for consonant identification might be dependent on the stimulation rate.

Why is it that 8 steps, and not 16 or 32 steps, are sufficient for accurate consonant recognition? The answer can be found by examining closely the quantized waveforms in Fig. 5. As one might expect, quantization affects the coding of temporal-envelope cues. These cues become progressively less clear as the number of steps is reduced from 32 to 2 (Fig. 5). The envelopes obtained using two and four steps provide very little, and possibly not perceptually meaningful, information about manner of articulation. That is, the patients will most likely perceive the presence or absence of the signal, but would not be able to tell whether the signal is a consonant or a vowel, let alone identify the consonant or vowel. To visual inspection, manner cues improve significantly when the number of steps is increased to eight. With eight steps, the voiced segment (the vowel /a/, in our case) of the word /apa/ can be easily discriminated from the unvoiced segment (the closure, burst and aspiration of /p/, in our case). This information about manner of articulation contained in each channel coupled with the across-channel envelope information allowed the implant listeners to identify the consonants accurately with only eight steps in the dynamic range.

The outcome that eight steps were enough for consonant recognition demonstrates that fine temporal-envelope cues are not needed for accurate consonant recognition when there are enough spectral cues. As shown in Fig. 5, the waveform that was quantized into eight steps lacks the temporal fine structure present in the original signal. Yet, the 8-step consonants were identified with roughly the same accuracy as the 512-step consonants. Drullman (1995) has also demonstrated that fine temporal cues are not needed when full spectral cues are available. He processed speech into 24 $\frac{1}{4}$-octave channels, computed the envelopes of each channel, and modulated the envelopes with noise. The processed waveforms had the original speech envelopes but the temporal structure of noise. High speech intelligibility was obtained even when the envelopes were quantized into two levels. In his study, however, the normal-hearing listeners had more spectral information (24 channels) compared to the cochlear-implant listeners (at most 6 channels) of this study. Two steps were not enough, in our study, to maintain high consonant recognition. We suspect that a greater number (eight) of steps was needed to compensate for the lower spectral resolution. This hypothesis is investigated in experiment 2.

Contrary to the Zeng and Galvin (1999) study, we found that performance drops significantly when the number of steps is reduced to two (binary representation). We suspect that the difference in outcome is due to two factors. First, in the Zeng and Galvin study, the two output levels used in the binary condition were both above threshold, whereas in our study, only one of the two levels was above threshold. Second, the speech-processing strategies, SPEAK and CIS, that were used in the two studies were different. The SPEAK strategy uses a large number of electrodes (22) and extracts

mostly spectral peak information, which is well preserved even for the binary representation of electrical amplitudes (see Fig. 3 in Zeng and Galvin, 1999). On the other hand, the CIS strategy of this study uses fewer number (six) of electrodes and extracts primarily temporal-envelope information from the acoustic signal. The results of the two studies taken together suggest that the number of intensity steps needed for speech recognition might be dependent on the number input channels supported by the cochlear implant processor. This hypothesis is investigated in experiment 2.

Experiment 1 demonstrated that only a handful of steps are needed by cochlear implant listeners for consonant identification under quiet conditions. However, under more realistic conditions with background noise, a greater number of steps might be needed. Zeng and Galvin (1999) noted that although the dynamic range reduction degraded phoneme recognition in quiet marginally, it had a significant effect in noise. Similarly, Wilson and colleagues (unpublished) noted that more steps were required to reach asymptotic performance in noise. More studies are needed to investigate the effect of number of steps on speech understanding under noisy conditions.

It should be noted that the amplitude quantization steps (Δi) of this study do not represent the discriminable steps measured by Nelson *et al.* (1996) using psychophysical methods. For some subjects in the Nelson *et al.* study, the amplitude DLs were not constant across the dynamic range but decreased as a function of the dynamic range, i.e., the step sizes were smaller in the upper dynamic range and relatively larger in the mid-to-low dynamic range. In our study, the step sizes (Δi) were uniform across the dynamic range, much like the amplitude DLs in two of the subjects in the Nelson *et al.* study. It is important to note, however, that the dynamic range was defined differently in the two studies. In our case, the dynamic range was defined from threshold to most-comfortable level (MCL), as commonly implemented in commercial speech processors, while in the Nelson *et al.* study it was defined from threshold to maximum acceptable loudness (MAL) level. If we exclude the upper dynamic range, from MCL to MAL, where the step sizes are smaller, we are left with the mid-to-low dynamic range where the step sizes are more uniform. On this view, the uniform quantization steps used in this study are a reasonable approximation to the uniform perceptual steps found in the low- to mid-dynamic range of implant users.

III. EXPERIMENT 2: ACOUSTIC HEARING

In experiment 1, we found that eight intensity steps within the dynamic range were enough for asymptotic consonant recognition by implant listeners who were using a six-channel CIS processor. In this experiment, we investigate whether this outcome holds when speech is processed through a larger (or smaller) number of channels. We hypothesize that there is a trade-off between spectral resolution (number of spectral channels) available and the intensity resolution (number of steps) needed. This hypothesis was motivated by our previous work on the need for accurate intensity resolution for processors using a small number of channels (Dorman *et al.*, 1997; Loizou *et al.*, 1998).

To produce speech with varying degrees of spectral resolution, speech was filtered through 4–20 frequency bands, and synthesized as a linear combination of sine waves with amplitudes extracted from the envelopes of the bandpassed waveforms, and frequencies equal to the center frequencies of the bandpass filters. To produce speech with varying degrees of intensity resolution, we quantized the amplitudes of the sinewaves to 2–32 levels. The intelligibility of vowels, consonants, and sentences was assessed as a function of spectral resolution and as a function of intensity resolution in normal-hearing listeners.

A. Method

1. Subjects

Nine graduate students from Arizona State University served as subjects. All of the subjects were native speakers of American English and had normal hearing. The subjects were paid for their participation.

2. Speech material

The test material included sentences, consonants in /aCa/ context and vowels in /hVd/ context. The consonant test was a subset of the Iowa consonant test (Tyler *et al.*, 1987) and consisted of 16 consonants in /aCa/ environment spoken by a single male speaker. Five repetitions of each consonant were used in a blocked and randomized test sequence. The vowel material consisted of the vowels in the words ''heed, hid, hayed, head, had, hod, hud, hood, hoed, who'd, heard.'' Each word was produced by three women and three girls. The stimuli were drawn from a set used by Hillenbrand *et al.* (1995). All the stimuli were presented in a completely randomized test sequence.

The sentence material was from the TIMIT database (Garofolo *et al.*, 1993). A different set of 15 sentences was used for each condition. A total of 540 sentences were randomly selected from the TIMIT database from the DR3 (north midland) dialect region. The sentences were produced by an equal number of female and male speakers—one sentence per speaker. The 540 sentences were divided into 36 lists (for 6 channel conditions $\times$6 quantization conditions), with 15 sentences in each list. Fifteen sentences were used for the first channel condition, 15 different sentences were used for the second channel condition, etc. There were eight sentences spoken by eight different male speakers and seven sentences spoken by seven different female speakers within each list. Each sentence contained, on the average, 7 words, and the 15 sentences in each list contained, on the average, a total of 100 words.

3. Amplitude quantization

Critical to the quantization process is the determination of the quantization step sizes, which are, in turn, dependent on the amplitude dynamic range. The dynamic range itself is a function of the frequency band (e.g., Boothroyd *et al.*, 1994). We therefore determined the quantization step sizes for each band separately.

We filtered the speech material (vowels, consonants, and sentences) through n (4–16) bandpass filters, and estimated

the envelopes of the bandpassed waveforms in each frequency band. We then determined the envelope-amplitude dynamic range (i.e., the difference between the maximum and minimum amplitudes) of each channel by computing the envelope histograms of the speech material. The maximum envelope amplitude in each channel, denoted as $X^i_{\max}$ where i is the channel number, was chosen to include 99% of all amplitude counts in that channel. The minimum envelope amplitude ($X^i_{\min}$) was set 0.5 dB above the rms value of the noise floor. The $X^i_{\max}$ and $X^i_{\min}$ values were then used to estimate the quantization step size, Δi, of each channel as follows:

$$\Delta_i = \frac{X^i_{\max} - X^i_{\min}}{Q-1}, \quad i = 1,2,...,n,$$

where Q is the number of quantization levels or steps, and n is the number of channels. Note that each channel had a different value for $X^i_{\max}$ and $X^i_{\min}$ since the envelope dynamic range of each channel was different. Consequently, the step sizes Δi were different in each channel. The $X^i_{\max}$ and $X^i_{\min}$ values are analogous to the most-comfortable (MCL) and threshold values in cochlear implants.

4. Speech synthesis

After estimating the quantization step sizes for each frequency band, we processed the speech material as follows. Signals were first processed through a preemphasis filter (1200-Hz cutoff), with a 3-dB/oct roll-off, and then bandpassed into n frequency bands ($n=4,6,8,12,16$) using sixth-order Butterworth filters. Logarithmic filter spacing was used for $n<8$ and mel spacing was used for $n\geqslant 8$. The center frequencies and the 3-dB bandwidths of the filters can be found in Loizou *et al.* (1999). The envelopes of the signal were extracted by full-wave rectification, and low-pass filtering (second-order Butterworth) with a 400-Hz cutoff frequency. The envelope amplitudes were estimated by computing the root mean-square (rms) energy of the envelopes every 4 ms. The envelope amplitudes were then uniformly quantized to Q discrete levels ($Q=2,4,8,16,32$). Sine waves were generated with amplitudes equal to the quantized envelope amplitudes and with frequencies equal to the center frequencies of the bandpass filters. The phases of the sinusoids were estimated from the fast Fourier transform (FFT) of the speech segment (McAulay and Quatieri, 1986). The sinusoids of each band were finally summed and the level of the synthesized speech segment was adjusted to have the same rms value as the original speech segment. We found the normalization of the synthesized speech segments to be absolutely necessary to maintain high levels of speech intelligibility. Without this normalization the synthesized consonant segments could have the same rms levels as the vowel segments, thereby affecting the consonant-to-vowel (CV) ratios. Pilot experiments showed that this normalization can have a dramatic effect on performance for the two- and four-step conditions. Therefore, to avoid having improper vowel-to-consonant ratios as a factor that might potentially confound our results, we decided to normalize the synthesized speech segments in order to preserve the original (natural) consonant-to-vowel ratios.

We also processed the speech material through a simulation of the SPEAK strategy used in the Nucleus-22 device (McDermott *et al.*, 1992; Loizou, 1998). We did that so that we could compare our results with those reported by Zeng and Galvin (1999). The signal was processed into 20 channels the same way as before, except that only the six (out of 20) highest channel amplitudes were used for synthesis in each 4-ms segment. We adopted the filter spacing used in the SPEAK strategy (Seligman and McDermott, 1995). Note that this signal processing strategy is sometimes referred to as the 6-of-20 strategy, since only 6 out of the 20 amplitudes are used for stimulation in each cycle.

In addition to the quantized speech material, we also processed speech as described above but without quantizing the envelope amplitudes. We used this condition for comparative reasons and refer to it as the ''unquantized'' condition.

5. Procedure

The experiment was performed on a PC equipped with a Creative Labs SoundBlaster 16 soundcard. The subjects listened to the speech material via closed ear-cushion headphones at a comfortable level set by the subject. A graphical interface was used that allowed the subjects to select the vowel or consonant they heard using a mouse. For the sentence material, subjects were asked to type in as many words as they could understand.

Before each condition, subjects were given a practice session with examples of speech (vowels, consonants, or sentences) processed through the same number of channels and the same number of steps in that condition. None of the sentences used in the practice was used in the test. A sequential test order, starting with speech material processed through a large number of channels ($n=16$) and continuing to speech material processed through a small number of channels ($n=4$), was employed. We chose this sequential test design to give the subjects time to adapt to listening to the altered speech signals. The test order for the different number of steps in each channel condition was counterbalanced between subjects.

B. Results

The results, scored in percent correct for vowels and consonants and in percent words correct for sentences, are shown in Fig. 6. A two-factor (channels and steps) repeated measures analysis of variance (ANOVA) on the sentence data showed a significant main effect of number of channels [$F(5,45)=371.6$, $p<0.0005$], a significant effect of number of steps [$F(5,45)=586.9$, $p<0.0005$], and a significant interaction between number of channels and number of steps [$F(25,225)=34.1$, $p<0.0005$]. *Post hoc* analysis, according to Fisher's LSD, showed that asymptotic performance for the 4-channel condition was reached with 16 steps, for the 6-, 8- and 12-channel conditions with 8 steps, and for the 16-channel and 6-of-20 channel conditions with 4 steps.

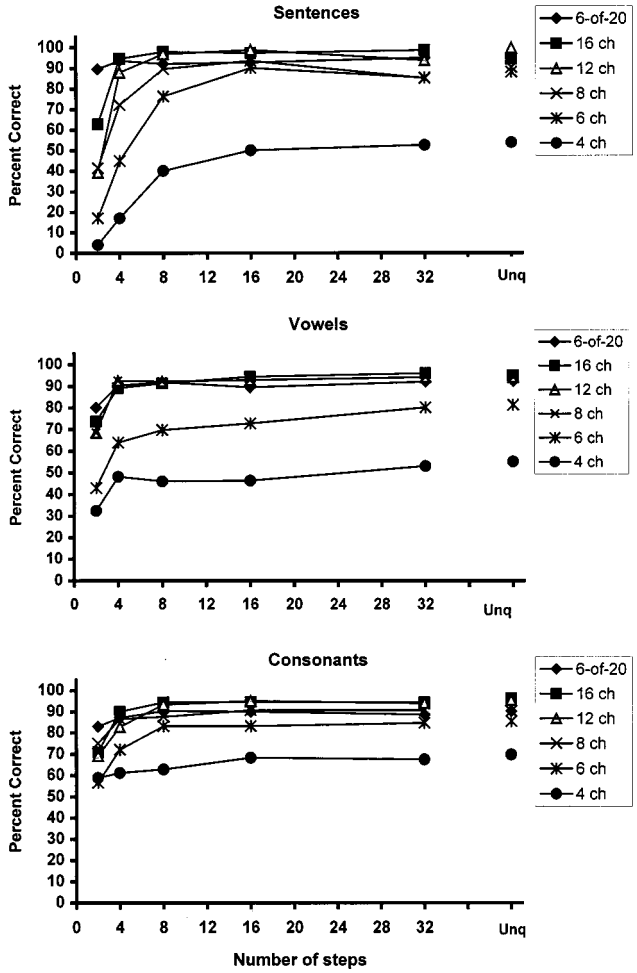


FIG. 6. Performance of normal-hearing listeners on sentence, consonant, and vowel recognition as a function of the number of channels and as a function of the number of steps. In the ''Unq'' condition, the speech materials were left unquantized.

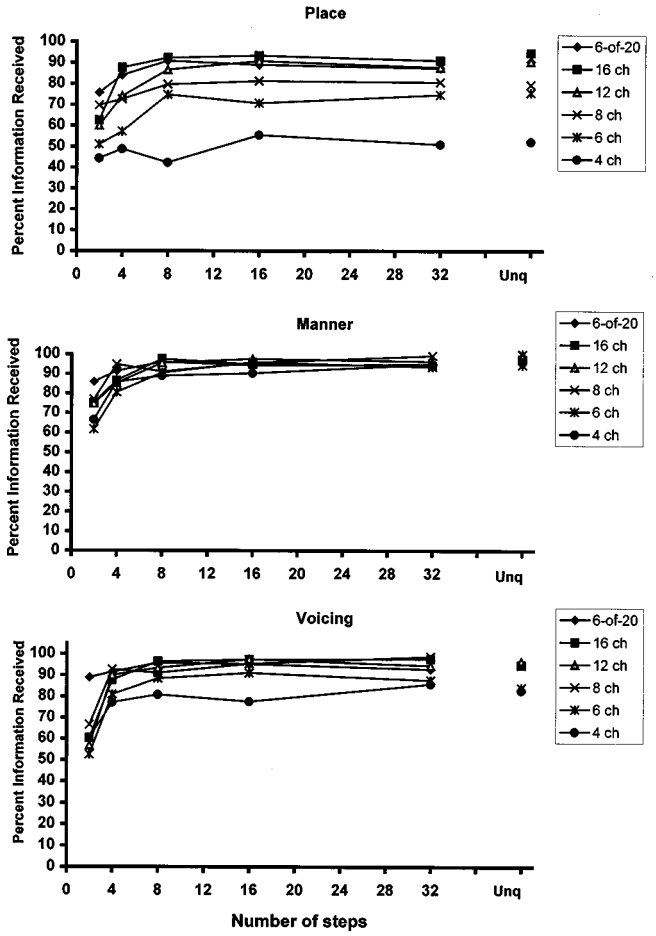


FIG. 7. Percent information received for the features ''place,'' ''manner,'' and ''voicing'' as a function of the number of channels and as a function of the number of steps.

A two-factor (channels and steps) repeated measures ANOVA on the vowel data showed a significant main effect of number of channels [$F(5,45)=100.3$, $p<0.0005$], a significant effect of number of steps [$F(5,45)=82.1$, $p<0.0005$], and a significant interaction between number of channels and number of steps [$F(25,225)=4.5$, $p<0.0005$]. *Post hoc* analysis, according to Fisher's LSD, showed that asymptotic performance for all channel conditions, except for the six-channel condition, was reached with four steps. For the six-channel condition, the vowel performance with 8 steps did not differ significantly ($p>0.05$) from the 4-, 16-, 32-step or the unquantized conditions.

Similar repeated measures ANOVA on the consonant data showed a significant main effect of number of channels [$F(5,45)=29.4$, $p<0.0005$], a significant effect of number of steps [$F(5,45)=112.3$, $p<0.0005$], and a significant interaction between number of channels and number of steps [$F(25,225)=8.3$, $p<0.0005$]. Post hoc analysis, according to Fisher's LSD, showed that asymptotic performance for the 6-, 12-, and 16-channel conditions was reached with eight steps. Consonant performance for the four-channel condition was low and uniform across the number of steps used. The score obtained with 2 steps did not differ significantly ($p=0.068$) with the 32-step and unquantized conditions. Asymptotic performance for the 8-channel and 6-of-20 channel conditions was reached with four steps.

The results for feature analysis (Miller and Nicely, 1955) for the consonants are shown in Fig. 7. A two-factor (channels and steps) repeated measures ANOVA performed on the three features (place, manner, and voicing) separately indicated a significant main effect of number of channels ($p<0.005$), a significant effect of number of steps ($p<0.005$) and a significant interaction ($p<0.005$) between number of channels and number of steps. For the feature ''place of articulation,'' *post hoc* analysis, according to Fisher's LSD, showed that asymptotic performance for the 6-, 8-, and 12-channel conditions was reached with eight steps, while for the 16-channel and 6-of-20 conditions it was reached with four steps. Performance for the four-channel condition was flat across the different numbers of steps used, in line with the percent-correct results. For the feature ''manner of articulation,'' *post hoc* analysis (Fisher's LSD) showed that asymptotic performance for the 6-, 12-, and 16-channel condition was reached with eight steps, while for the 4- and 8-channel conditions it was reached with four steps. Performance was flat for the 6-of-20 condition, with no statistically significant differences between the scores obtained with 2 and 32 steps. For the feature ''voicing,'' *post hoc*

analysis (Fisher's LSD) showed that asymptotic performance for the 6-, 8-, 12-, and 16-channel conditions was reached with only four steps. Performance was flat for the 4-channel and 6-of-20 conditions.

C. Discussion

The statistical analysis indicated, consistent with our original hypothesis, a significant interaction between spectral resolution and intensity resolution for all test materials. The minimal number of steps needed to reach asymptotic performance was a function of the number of channels available and the speech material used. For the sentence test, the number of steps needed to obtain high levels of intelligibility was found to be inversely proportional to the number of channels available. Sixteen steps were needed to reach asymptotic performance for speech processed through four channels. In contrast, four steps were sufficient for speech processed through 16 or more channels. High sentence intelligibility scores (90%) were obtained even with two steps when speech was processed through the 6-of-20 strategy. The sentence results confirm our original hypothesis about the tradeoff relationship between spectral resolution and intensity resolution. When the spectral resolution is poor (i.e., a small number of channels is available) relatively fine intensity resolution is needed to achieve a high level of speech understanding. Conversely, when the spectral resolution is good (i.e., a large number of channels is available) fine intensity resolution is not needed to achieve high levels of understanding.

A different pattern of performance was obtained for vowels. Asymptotic performance was obtained with four steps in all channel conditions except for $n=6$. Vowel performance seemed to be more dependent on the number of channels available, which was not surprising since fine spectral resolution is needed for accurate vowel identification. High vowel identification scores ($>90\%$) were obtained with eight or more channels, in agreement with our previous findings (Dorman *et al.*, 1997). The vowel performance obtained with eight or more channels and four steps was not significantly different than the vowel performance obtained in the unquantized condition.

Compared to vowels, a greater number of steps as needed to identify consonants. Eight steps were needed for most channel conditions to reach asymptotic performance. Eight steps were needed for consonants processed through six channels, consistent with our findings with cochlear implant listeners in experiment 1. The finding that more steps are needed to identify consonants than vowels is not surprising, since the quantization affects the coding of temporal-envelope cues, which are more important for consonant identification than vowel identification. Feature analysis also indicated that eight steps are needed for the features ''place'' and ''manner'' to reach asymptotic performance in most channel conditions, while only four steps are needed for the feature ''voicing.'' Overall, eight steps are needed to reliably transmit information about place, manner, and voicing, consistent with the number of steps needed to reach asymptotic performance on consonant recognition.

The data from the Zeng and Galvin (1999) study are entirely consistent with the data from normal-hearing listeners (experiment 2) if we assume that the Nucleus-22 implant listeners have 20 functional channels. However, several experiments have shown that this is not the case. Fishman *et al.* (1997) and Wilson (1997) have demonstrated that the performance of N22 implant listeners reaches asymptote when only four to seven channels are activated. This also seems to be the case with the N22 implant listeners in the Zeng and Galvin study. For vowels, the N22 patients' level of performance at asymptote (20 electrodes activated and full dynamic range) was approximately 78% correct, and for consonants it was approximately 70% correct, which is equivalent to the performance of normal-hearing listeners using four to six channels of stimulation (Fig. 6). Furthermore, increasing the number of electrodes activated from 4 to 20 improved the consonant score by only 10% (50% versus 60% correct), whereas in our study (Fig. 6) the mean consonant score improved by approximately 40% (60% versus 100% correct) when the number of channels increased from 4 to 20. This suggests that, most likely, the N22 patients were not using 20 channels of information when all 20 electrodes were activated, but rather 4 to 6 channels. Thus there is converging evidence that (*i*) the N22 patients receive the equivalent of 4 to 6 channels of information, and (*ii*) the N22 patients are not performing at the level of a 6-of-20 processor. If we then assume that the N22 patients were using only four channels, then the pattern of performance of our vowel and consonant data (Fig. 6) falls in line with the data in the Zeng and Galvin study. With four channels, there is hardly any difference in vowel or consonant performance as we increase the number of steps from 2 to 32 (Fig. 6), much like the data in the Zeng and Galvin study on vowel and consonant recognition.

It is clear from our studies on dynamic range reduction and number of channels that degrading amplitude information and degrading frequency information have different effects on sentences, as compared to consonants and vowels. Degrading frequency information (i.e., reducing number of channels) has a large effect on consonants and vowels but a smaller effect on sentences (Dorman *et al.*, 1997; Shannon *et al.*, 1995). In contrast, degrading amplitude information has a large effect on sentences but a smaller effect on consonants and vowels. The latter outcome may be the result of the loss of amplitude-based word and syllabic segmentation cues in the time waveform (see Zue, 1985).

IV. CONCLUSIONS

(1) For five cochlear implant listeners using a six-channel CIS processor, eight quantized amplitude steps were sufficient for achieving asymptotic levels of consonant recognition in quiet. Eight steps were sufficient for both better performing and poorly performing patients. These results suggest that amplitude resolution is not a major factor in determining consonant identification.

(2) The outcome that eight steps were enough for consonant recognition provides yet another demonstration (e.g., Drullman, 1995) that fine temporal-envelope cues are not needed for accurate consonant recognition when

there are enough spectral cues (a maximum of six channels, in our case, for the implant listeners).

(3) The implication of these findings for the design of cochlear implant processors is that the compression function, which is used to map the acoustic input to electrical output, need not be defined with high resolution. Consequently, the size of the compression tables stored in processor memory could be greatly reduced.

(4) For normal-hearing listeners listening to speech processed through 4–20 channels and channel amplitudes quantized to 2–32 steps, the number of steps needed to achieve asymptotic performance was a function of the number of channels and the speech material used.

(5) For vowels, asymptotic performance was obtained with four steps and was independent of the number of channels. For consonants, eight steps were needed for most channel conditions to reach asymptotic performance. This is consistent with our findings with cochlear implant listeners in experiment 1.

(6) For sentences 16 steps were needed to reach asymptotic performance with 4 channels, while 4 steps were needed for sentences processed through 16 or more channels. High sentence intelligibility scores (90%) were obtained even with two steps when speech was processed through a 6-of-20 strategy (SPEAK type). These results reaffirm our hypothesis that there is an inverse relationship between spectral resolution and intensity resolution. When spectral resolution is poor (i.e., a small number of channels is available) relatively fine intensity resolution is needed to achieve high levels of understanding. Conversely, when the intensity resolution is poor, a high degree of spectral resolution is needed to achieve asymptotic performance.

(7) The results of this study, taken together with previous findings (Zeng and Galvin, 1999; Loizou *et al.*, 2000) on the effect of reduced dynamic range, suggest that the performance of cochlear implant subjects is primarily limited by the small number (four to six) of channels received, and not by the small number of steps or reduced dynamic range. Therefore, more research is needed to find ways to increase the number of channels transmitted by cochlear implants.

ACKNOWLEDGMENTS

The authors would like to thank Fan-Gang Zeng and the anonymous reviewers for providing valuable suggestions. This research was supported by Grant Nos. R01 DC03421 and R01 DC00654 from the National Institute of Deafness and other Communication Disorders, NIH.

[1]The range in number of discriminable steps (8–45 steps) estimated in the Nelson *et al.* (1996) study is actually an overestimate of the true number of steps that an implant listener might have with a commercial speech processor. This is because Nelson *et al.* calculated the number of DLs over the dynamic range between threshold and maximum acceptable loudness (MAL) levels. In commercial speech processors, however, speech is typically mapped between threshold and most comfortable level (MCL), which is smaller than the MAL level. And, since the smallest DLs were found between MCL and MAL levels, the number of discriminable steps associated with dynamic range in speech processors may be significantly smaller than those estimated by Nelson *et al.* (1996).

[2]The power exponent p was set equal to -0.0001 to match the logarithmic compression function, of the form $y = A \log(1+cx)+B$, used in the Med-El device, where A and B are constants used for mapping the acoustic signal x between threshold and most comfortable level, and c is a constant that controls the shape of the compression function ($c=512$, in our case). It should be noted that the value of the power exponent p corresponding to logarithmic mapping depends on the input range, and, in particular, on the minimum value of x ($X_{\min}$). If $X_{\min}=0$, then the value of $p=0.2$ will yield a logarithmic mapping similar to the one used in the Med-El device, whereas if $X_{\min}=1$, then the value of $p=-0.0001$ will yield a logarithmic mapping. In our implementation, X min$=1$.

Boothroyd, A., Erikson, F., and Medwetsky, L. (**1994**). ''The hearing aid input: A phonemic approach to assessing the spectral distribution of speech,'' Ear Hear. **15**, 432–442.

Dillier, N., Spillman, T., and Guntensperger, J. (**1983**). ''Computerized testing of signal encoding strategies with round window implants, in *Cochlear Prostheses: An International Symposium*, edited by C. W. Parkins and S. W. Anderson, Ann. N.Y. Acad. Sci. **405**, 360–369.

Dorman, M., Loizou, P., and Rainey, D. (**1997**). ''Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs,'' J. Acoust. Soc. Am. **102**, 2403–2411.

Drullman, R. (**1995**). ''Temporal envelope and fine structure cues for speech intelligibility,'' J. Acoust. Soc. Am. **97**, 585–592.

Fishman, K., Shannon, R., and Slattery, W. (**1997**). ''Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant processor,'' J. Speech Hear. Res. **40**(5), 1201–1215.

Francois, J., Tinembart, J., Bessat, C., Leone, P., Rossman, F., and Pelizzone, M. (**1994**). ''Implants cochle aires: Un processeur portable pour le developpement de l'algorithme CIS,'' *Actes de le conference DSP 94*, Paris.

Garofolo, J., Lamel, L., Fisher, W., Fiscus, J., Pallett, D., and Dahlgren, N. (**1993**). ''DARPA TIMIT: Acoustic-phonetic continuous speech corpus,'' *NIST Technical Report* (distributed with the TIMIT CD-ROM).

Hillenbrand, J., Getty, L., Clark, M., and Wheeler, K. (**1995**). ''Acoustic characteristics of American English vowels,'' J. Acoust. Soc. Am. **97**, 3099–3111.

Hochmair-Desoyer, I., Hochmair, E., and Fischer, R. (**1981**). ''Four years of experience with cochlear prostheses,'' Med. Prog. Technol. **8**, 107–119.

Loizou, P. (**1998**). ''Mimicking the human ear: An overview of signal processing techniques for converting sound to electrical signals in cochlear implants,'' IEEE Signal Process. Mag. **15**(5), 101–130.

Loizou, P., Dorman, M., and Fitzke, J. (**2000**). ''The effect of reduced dynamic range on speech understanding: Implications for patients with cochlear implants,'' Ear Hear. **21**(1), 25–31.

Loizou, P., Dorman, M., and Powell, V. (**1998**). ''The recognition of vowels produced by men, women, boys and girls by cochlear implant patients using a six-channel CIS processor,'' J. Acoust. Soc. Am. **103**, 1141–1149.

Loizou, P., Dorman, M., and Tu, Z. (**1999**). ''On the number of channels needed to understand speech,'' J. Acoust. Soc. Am. **106**, 2097–2103.

McAulay, R., and Quatieri, T. (**1986**). ''Speech analysis/synthesis based on a sinusoidal representation,'' IEEE Trans. Acoust., Speech, Signal Process. **ASSP-34**(4), 744–754.

McDermott, H., McKay, C., and Vandali, A. (**1992**). ''A new portable sound processor for the University of Melbourne/Nucleus Limited multielectrode cochlear implant,'' J. Acoust. Soc. Am. **91**, 3367–3371.

Miller, G., and Nicely, P. (**1955**). ''An analysis of perceptual confusions among some English consonants,'' J. Acoust. Soc. Am. **27**, 338–352.

Nelson, D., Schmitz, J., Donaldson, G., Viemeister, N., and Javel, E. (**1996**). ''Intensity discrimination as a function of stimulus level with electric stimulation.'' J. Acoust. Soc. Am. **100**, 2393–2414.

Pfingst, B., Burnett, P., and Sutton, D. (**1983**). ''Intensity discrimination with cochlear implants,'' J. Acoust. Soc. Am. **73**, 1283–1292.

Poroy, O., and Loizou, P. (**2000**). ''Development of a speech processor for laboratory experiments with cochlear implant patients,'' *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Instanbul, Turkey.

Schroder, A., Viemester, N., and Nelson, D. (**1994**). ''Intensity discrimination in normal-hearing and hearing-impaired listeners,'' J. Acoust. Soc. Am. **96**, 2683–2693.

Seligman, P., and McDermott, H. (**1995**). ''Architecture of the Spectra 22 speech processor,'' Ann. Otol. Rhinol. Laryngol. **Suppl. 166**, 139–141.

Shannon, R. (**1983**). ''Multichannel electrical stimulation of the auditory nerve in man. I. Basic Psychophysics,'' Hear. Res. **11**, 157–189.
Shannon, R., Jensvold, A., Padilla, M., Robert, M., and Wang, X. (**1999**). ''Consonant recordings for speech testing,'' J. Acoust. Soc. Am. **106**, L71–L74.
Shannon, R., Zeng, F., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). ''Speech recognition with primarily temporal cues,'' Science **270**, 303–304.
Tyler, R., Preece, J., and Lowder, M. (**1987**). The Iowa audiovisual speech perception laser videodisc, *Laser Videodisc and Laboratory Report*, Dept. of Otolaryngology, Head and Neck Surgery, University of Iowa Hospital and Clinics, Iowa City.
Wilson, B. (**1997**). ''The future of cochlear implants,'' Br. J. Audiol. **31**, 205–225.
Wilson, B., Finley, C., Lawson, D., Wolford, R., Eddington, D., and Rabinowitz, W. (**1991**). ''Better speech recognition with cochlear implants,'' Nature (London) **352**, 236–238.
Zeng, F-G., and Galvin, J. (**1999**). ''Amplitude mapping and phoneme recognition in cochlear implant listeners,'' Ear Hear. **20**, 60–74.
Zeng, F-G., and Shannon, R. (**1999**). ''Psychophysical laws revealed by electric hearing,'' NeuroReport **10**, 1931–1935.
Zue, V. (**1985**). ''The use of knowledge in automatic speech recognition,'' Proc. IEEE **73**, 1602–1615.

Air-wood coupling and the Swiss-cheese violin

Gabriel Weinreich, Colin Holmes, and Maureen Mellody
Randall Laboratory of Physics, The University of Michigan, Ann Arbor, Michigan 48109-1120

(Received 24 March 2000; accepted for publication 9 August 2000)

Some problems with the conventional formalism for describing the coupling of fluid vibrations to those of an enclosing shell are examined. An alternative (''dynamic'') basis for expanding the normal modes, in which the ''pure shell modes'' include incompressible motion of the fluid, is proposed. This new approach is applied to Hutchins's ''Swiss-cheese violin,'' the behavior of whose air modes for the case of a rigid shell have been calculated by Shaw. Measurements are presented of various response functions of this instrument as a function of both frequency and the number of open rib holes. These results show the predicted ''mode veering'' behavior, and can be easily matched to theory with the assumption of plausible parameters. © *2000 Acoustical Society of America.* [S0001-4966(00)02611-4]

PACS numbers: 43.75.De, 43.20.Ks, 43.40.At [RDA]

I. DYNAMICS OF COUPLED MODES

A. A Word about nomenclature

Occasionally, one hears the word ''mode'' used synonymously with ''normal mode'' (or ''eigenmode''), designating a pattern of motion which, once started in a given (linear) system, continues to oscillate sinusoidally in time without changing its shape.[1] With such an interpretation, the phrase ''coupled modes'' becomes a contradiction in terms, since two vibration patterns which can each persist independently are, by definition, uncoupled.

For this reason, it is important to point out that our use of ''mode'' in this paper is considerably broader. In particular, we note that the set of motions that comprise the normal modes of a system form an *orthogonal* and *complete* set; that is, any pattern of displacement of the same coordinates can be expressed as a linear combination of the normal modes. Conversely, any actual pattern of displacement of such a system can, at any instant, be expressed as a linear combination of the normal modes of a hypothetical system having the same coordinates but different dynamics (as defined, for example, by its Lagrangian). We then refer to this set of ''hypothetically normal modes'' as a ''basis set'' or simply ''basis,'' and do not hesitate to use the word ''mode'' for any of its members.

Such a nomenclature is especially useful if the hypothetical system, rather than being completely different from the one under study, is in fact a closely related one. In that case we can say that, in the actual system, the normal modes of the hypothetical one have been *coupled*. The amount of this coupling can be quantified by examining to what degree the coefficients of expansion of the actual normal modes in terms of the hypothetical basis differ from a simple Kronecker delta, and relating this difference to the discrepancy between the two Lagrangians.

B. Description of the basis

It is surely a truism that among the normal modes of a violin there is really no such thing as a ''pure air mode'' or ''pure wood mode''; rather, every normal mode involves motion of both those subsystems. Yet it is just as obvious that a careful description of what we call ''coupled modes'' requires, first, a definition of what exactly the modes are which are to be coupled. As explained in the previous section, we need to define an orthogonal *basis set* of modes which need not themselves be solutions of the equations of motion of our system, but in which those solutions can be *expanded*; the resulting expansion coefficients then become the variables in terms of which the coupled motion can be quantitatively discussed. For the case of a vibrating enclosure such as we are dealing with here, one thinks first of taking as the basis a set of hypothetical ''pure wood modes,'' in which only the wood moves, and a set of hypothetical ''pure air modes,'' in which only the air moves.

Although it is mathematically possible to deal with such a basis,[2] it is decidedly unsatisfactory from the point of view of physical intuition. The reason is that no system can be imagined that we can think of as the *uncoupled version* of the one we are interested in, in the sense of having the above ''pure'' modes as eigenmodes. Specifically, since the wooden shell acts as the boundary condition for the air motion, *it is kinematically impossible for the wood to be moving while the air remains everywhere stationary, quite apart from any assumed dynamics*. The traditional discussion, which evades this problem by defining the ''pure wood modes'' as taking place after the system has been placed in a vacuum, is unacceptable, in that a valid basis set must consist of prescribed motions for all, and not merely some, of the system's degrees of freedom—even if for some of them the ''prescribed motion'' consists of being at rest.

We should note that it *is* possible to construct in our minds a wood-air system in which the motion of the wood is *dynamically* independent of the presence of air. The criterion for this is that $z \equiv (B\rho_0)^{1/2}$, which makes the characteristic impedance of the air negligible; here B is the adiabatic bulk modulus of the air, and ρ_0 its density. One possibility is thus to go back to the vacuum condition but, instead of removing the air altogether, make its pressure very small. If this is done isothermally, B and ρ_0 will go down in proportion, and consequently so will z, while the speed of sound c

$=(B/\rho_0)^{1/2}$ will not change. It is now possible to solve for the ''pure wood modes'' as before, then use those solutions as given moving boundary conditions and solve for the air motion to be associated with each mode. But this is a hard calculation: since, for most modes, λ_{air} is smaller than the dimensions of the cavity, solving for the air motion entails solving a true internal wave problem every time. In addition, any change of parameters which affects the frequency of the pure wood mode will change λ_{air}, and will thus require a recalculation of the air motion.

That is why, in this paper, we propose a different approach: in defining the basis, we assume that B remains constant while ρ_0 becomes vanishingly small, so that z becomes negligible while c moves toward very large values. This again guarantees that the forces which the air exerts on the wood can be neglected; but, since now λ_{air} for each mode grows without limit, *the motion of the air becomes incompressible*. Thus the solution for the air motion to be associated with a wood mode, while perhaps still not quite trivial, becomes one of incompressible hydrodynamics rather than wave motion, and is no longer dependent on a characteristic length which can change. This approach also allows us to resolve the paradox of moving boundary conditions for the air modes: we define ''pure air modes'' as comprising air motion over and above the incompressible flow attached to what we call ''pure wood modes.'' For this incremental flow pattern, the boundary conditions are, indeed, stationary.

The final definition of our basis, which we distinguish from the conventional one by the adjective ''dynamic,'' accordingly makes each member either

(a) a ''*dynamic* pure wood mode'' (whose amplitude will be denoted by x_i), for which the air, rather than being absent, moves incompressibly in response to the shell without exerting any forces on it; in other words, if the shell flexes so as to decrease its volume, an equal amount of air flows out through the *f*-holes (as would be the case with a tube of toothpaste); or

(b) a ''*dynamic* pure air mode'' (whose amplitude will be denoted by ξ_σ), which describes the motion of the air when the wood is constrained to immobility; equivalently, it describes the *deviation* of air motion from the incompressible pattern entailed by a dynamic pure wood mode. Like dynamic pure wood modes, dynamic pure air modes become normal modes if the density of the air is made to vanish at finite pressure—but the corresponding normal frequencies in that case become very large.

The density of an ideal gas can be decreased at constant bulk modulus either by decreasing its molecular weight or by increasing its temperature. Although neither process can be carried indefinitely far in practice, this does not disqualify the thought experiments required to understand the coupling.

C. Equations of motion

In terms of the dynamic pure modes, the potential energy is simply a sum of squares,

$$V=\Sigma_i \tfrac{1}{2}k_i x_i^2+\Sigma_\sigma \tfrac{1}{2}\kappa_\sigma \xi_\sigma^2. \tag{1}$$

This distinguishes the ''dynamic'' from the ''conventional'' formulation, since conventionally the enclosed air is compressed both by the shell when it moves so as to decrease the volume, and by the influx of additional air. In our ''dynamic'' formulation, however, the equivalent incompressible air flow is already built into x_i; hence, whether the wood is moving or not, the actual compression of the air comes entirely from the ξ_σ.

Unlike the potential energy, which is a sum of squares, it is now the kinetic energy that must contain the coupling terms. They arise from the fact that, at any point in space, the air velocity is a superposition of the incompressible flow associated with the (dynamic) wood mode and the flow velocity of the (dynamic) air mode itself; and when this is squared to obtain the kinetic energy, cross terms between the mode amplitudes appear.

The correct parametrization of the total kinetic energy will involve overlap integrals of the air flow patterns of all the various modes. For illustrative purposes, however, we shall here simplify the picture considerably by invoking the ''quasispherical'' model of the Appendix (Sec. A 2), according to which there is only one ''important'' air mode lying within the monopole regime. This means that there is only one ξ_σ, denoted simply by ξ. In addition, the nature of this mode (the so-called $A0$, Sec. A 2) is such that the kinetic energy density is concentrated in the immediate vicinity of the *f*-holes.

In this approximation, the total kinetic energy takes the rather simple form

$$T=\Sigma_i \tfrac{1}{2}m_i \dot{x}_i^2+\tfrac{1}{2}\mu(\dot{\xi}+\Sigma_j \beta_j \dot{x}_j)^2, \tag{2}$$

where the first term is the kinetic energy of the wood, and the second is the kinetic energy of the air. Concretely, we define ξ as *the volume of air passing through the f-holes* in the single air mode (for which, of course, the wood is stationary), thus determining the numerical value of the constant κ. The constant μ is the appropriate inertia coefficient, and the β_j define effective volume changes of the shell per unit x_j. The exact scaling of the wood coordinates and their associated stiffnesses are left open.

By now constructing the Lagrangian as $T-V$ and applying the Lagrangian equations, we obtain the equations of motion of our system. Since our interest is in normal modes, rather than in any transient behavior, we can also replace the second time derivative by the factor $-\omega^2$. The result is then the following set of equations:

$$(k_i-\omega^2 m_i)x_i-\omega^2\mu\beta_i\Sigma_j\beta_j x_j-\omega^2\mu\beta_i\xi=0, \quad i=1,2,..., \tag{3}$$

$$-\omega^2\mu\Sigma_j\beta_j x_j+(\kappa-\omega^2\mu)\xi=0. \tag{4}$$

As expected in this formulation, each of the variables (x_i,ξ) is now governed by its own independent equation, *except for coupling terms that are proportional to* μ.

It should be noted that, had we used the ''conventional'' rather than the ''dynamic'' basis, the kinetic energy would be a sum of squares, with all coupling terms arising from the potential energy. The identification of a ''coupling constant'' would then be considerably more difficult.

D. Sound hole sum rule and radiativity of dynamic basis

The sound hole sum rule,[3] which requires the monopole radiativity of the violin (or any other instrument having a flexible shell and one or more sound holes) to vanish as ω^2 at low frequencies, follows from the fact that, in a system of limited overall size, air behaves more and more incompressibly as the wavelength is increased. Hence, in the limit of low frequencies, any change in shell volume is precisely compensated by air motion through the *f*-holes, so that the overall volume cannot change and the monopole moment becomes exactly zero.

The same principle applies much more powerfully to the ''dynamic basis'' that we have adopted. First of all, we note that a dynamic pure wood mode includes incompressible air motion by definition, so again it is not capable of developing any monopole moment. At the same time, as we saw, the limit of vanishing air density means that the speed of sound becomes very large, so that the monopole limit—the frequency below which monopole radiation predominates—itself tends toward infinity (Appendix, Sec. A 1).

We are thus led to formulate two rather startling new principles:

(1) Under conditions that define the dynamic basis, *pure (dynamic!) wood modes are incapable of radiating sound*. Any sound radiation must come from the (dynamic!) air modes.
(2) Since the energy of vibration comes ultimately from the vibrating strings, and the strings can only be connected to the wood but not directly to the air, it follows that *the only way in which a violin can produce any sound at all is through the mechanism of (dynamic!) air-wood coupling*. This principle forms the central motivation of the present paper.

II. SWISS-CHEESE VIOLIN

A. Description of the instrument

The ''Swiss-cheese violin'' (SUS No. 185), playfully dubbed ''Le Gruyère'' by its designers, was built by Carleen Hutchins in 1982 in consultation with E. A. G. Shaw and A. H. Benade.[4–6] It is a violin of otherwise standard design in whose ribs a number of holes have been drilled in such a pattern as to minimize, as much as possible, their effect on the mechanical properties of the instrument and hence on the nature of its wood dynamics—though, naturally, their effect on the air dynamics is profound. By plugging the holes with small corks, the behavior of the air can be returned more or less to what it would be in a standard instrument. Also available is a full set of open-cell foam plugs which were not, however, utilized in the work reported here.

In the original papers describing this instrument,[4,5] the 65 rib holes were specified as having a diameter of 5.5 mm. Our own examination revealed, however, that their diameters ranged from a minimum of 5.81 mm to a maximum of 8.32 mm; in addition, there are a number of even larger holes which had been plugged semi-permanently with modeling clay, and which we left that way. (We believe that their original purpose was to insert microphones into the cavity.) Ignoring the holes in the *C*-bouts (since we did not open them in our main data sets, see Sec. IV B) narrows the diameter range of the 49 remaining holes to 6.11–8.32 mm, with an average value of 6.26 mm.

To understand the relevance of Le Gruyère to the discussion of this paper, we need to return to the parameter μ introduced in Sec. I C, where it was described as ''the appropriate inertia coefficient'' to be used in converting the square of $\dot{\xi}$ to a kinetic energy, ξ having been defined as the volume of air passing through the openings in the shell. Now if an otherwise closed shell is given N identical holes, and if the rate of volume change of the internal air is $\dot{\xi}$, then the volume rate of flow through each hole will be $\dot{\xi}/N$; assuming that the flow patterns through the various holes do not overlap, this will give a kinetic energy associated with each hole of

$$T_1 = \tfrac{1}{2}L(\dot{\xi}/N)^2, \tag{5}$$

where L is the so-called *inertance*, or acoustical inductance, of the hole.[7] The total kinetic energy of the air flow will then be

$$T = NT_1 = \tfrac{1}{2}(L/N)\dot{\xi}^2. \tag{6}$$

Comparing this with Eq. (2), we find that, for this case,

$$\mu = L/N. \tag{7}$$

In actual fact, of course, the violin shell is not ''otherwise closed,'' as the previous paragraph assumed, because of the open *f*-holes. If we continue to use L for the inertance of a single open ''Swiss-cheese'' hole, then the inertance L_f of the pair of *f*-holes can be expressed in terms of an equivalent number N_f of such smaller holes by defining

$$N_f \equiv L/L_f. \tag{8}$$

The total kinetic energy then becomes

$$T = \tfrac{1}{2}L(N_f + N)^{-1}\dot{\xi}^2. \tag{9}$$

In Sec. I B, we defined the basis of pure wood modes and pure air modes as comprising the normal modes of the system in the limit as ρ_0, the density of the air, becomes small. Now, Eq. (9) suggests that the same regime can be approached by opening as many corks as possible in the Swiss-cheese violin; in fact, the reciprocal of the equivalent μ is linear in the number of corks that are opened. Thus Le Gruyère becomes a valuable tool in investigating the process by which the wood and air modes of a violin are coupled.

One word of caution is required: although opening a lot of holes in the ribs is equivalent to decreasing μ, the speed of sound c does not change in the process, as it would if ρ_0 were varied. Accordingly, the monopole limit of the violin, rather than increasing drastically, remains exactly the same, namely in the vicinity of 850 Hz,[8] and the conclusion of Sec. I D that, if the modes are decoupled by diminishing the air density, the violin will become totally incapable of sound radiation, does not apply. The violin would, however, truly become incapable of *monopole* sound radiation if the lowest air mode were to move well above the monopole limit.

B. Shaw's calculations

In a beautiful paper,[6] Shaw computes the frequencies of the lowest two air modes of Le Gruyère as a function of the number of rib holes that are open. He does this by representing the air in the instrument as a lumped-constant network consisting of two capacitances corresponding to the springiness of the upper and lower bouts, a series inertance for the constriction between the two *C*-bouts, and inertances to ground for the open *f*-holes as well as any additional rib holes from which the corks have been removed. The numerical values of the parameters are first estimated by Shaw *a priori*, then corrected to fit the measured[5] frequencies and pressure ratios of the A0 and A1 modes of this violin with the rib holes closed. The inertance of each hole, to be inserted into the equivalent circuit when a cork is opened, is computed *a priori* from the thickness of the ribs and the diameter of the hole. In calculating the frequencies with some rib holes open, Shaw assumes, for concreteness, that there are always twice as many holes open in the lower (larger) bout as in the upper (smaller) bout.

Of course, the use of a lumped-constant model does exact its price. In particular, Shaw places the inertances that correspond to open rib holes at the two ends of his circuit (see his Fig. 9), as though they had all been drilled either at the bottom of the violin, near the button, or at the top, near the neck. This is a reasonable way of representing the system if one thinks of the cavity as consisting of two sub-cavities connected by the constriction of the *C*-bouts (as Shaw's model does), in which case the pressure amplitude is always constant throughout either one of those bouts. In fact, however, there are important deviations from this behavior, as our data will show (Sec. IV B).[9] One should also remember that, in addition to appreciable variations in pressure both in the lower and in the upper bouts, a certain number of rib holes are located within the *C*-bouts themselves (where the variation is even greater, radically so for the A1).

C. Previous measurements

Data for Le Gruyère were reported first by Hutchins.[5] Of her various measurements, the ones that throw the most direct light on the air vibration are those taken by driving the system with a miniature loudspeaker inside the violin, and recording the response of a microphone which is also on the inside. Hutchins provided data on two cases—six corks removed, and all corks removed—which, as far as one could tell, fit well with Shaw's calculations.

A few years later, Rodgers[10] reported measurements on Le Gruyère by a completely different method: a spectrum obtained from the sound of the violin when a *glissando* is executed on each of the four strings. Although the frequencies obtained by Rodgers agree with those of Hutchins (and the calculated ones of Shaw) where they can be compared, many of the situations investigated by him do not correspond to the 1:2 ratio of corks removed from the upper/lower bout that was used for Shaw's calculation, so no direct comparison is possible. In addition, the method used by Rodgers for identifying the ''air modes'' is not definitive.

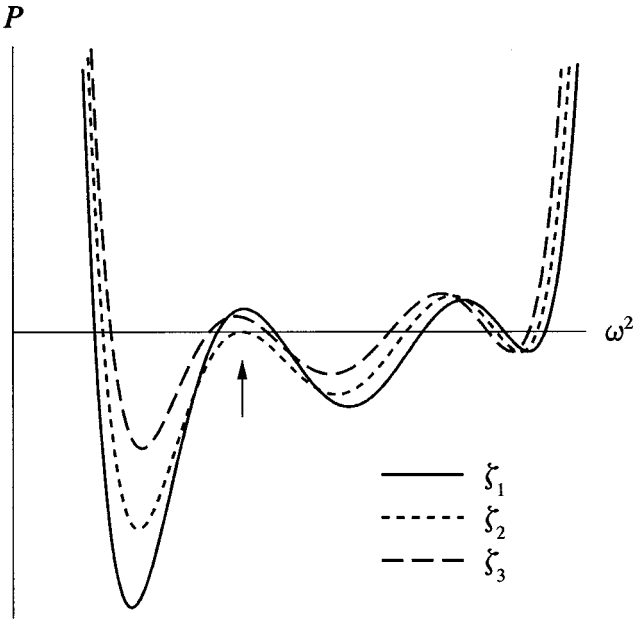


FIG. 1. Generic behavior of the polynomial $P_\zeta(\omega^2)$ for three values of ζ. The arrow points to a coalescing of two eigenvalues.

D. Algebra of crossing modes

As is well known, if a parameter ζ of a system is varied in such a way that the order of frequencies of two modes is reversed, in general no actual crossing occurs. Instead, as their frequencies approach, the modes themselves continuously change into one another, at the same time interchanging the paths that their frequencies had been following as a function of ζ. This behavior is aptly referred to as ''mode veering.'' [11]

The scale on which the veering takes place is determined by some appropriate ''coupling constant'' K which has the same dimensions as ζ, defined so that for $K=0$ the two mode frequencies do cross; let $\zeta=\zeta_0$ be the point where they do so. For $K\neq 0$, the veering is then confined to a region of approximate size K around ζ_0. In other words, when $|\zeta-\zeta_0|>K$, the frequencies of the two modes vary more or less the way they would in the absence of coupling; but when $|\zeta-\zeta_0|<K$ they veer so as to change places, while at the same time the vibration patterns associated with those modes change into each other as well.

To see the reason for veering behavior, we recall that a system of homogeneous linear equations can have a solution only if the determinant of coefficients vanishes. For equations such as (3) and (4), the determinant is a polynomial $P_\zeta(\omega^2)$ in ω^2 whose own coefficients are functions of ζ. When this polynomial is plotted against ω^2, the eigenvalues will lie at the values of ω^2 where $P_\zeta(\omega^2)$ crosses the horizontal axis (Fig. 1). Two mode frequencies can coincide only if P is tangent to the axis (as indicated by the arrow).

Now although it is indeed possible (depending on whatever other parameters determine the coefficients of the polynomial) for such a tangency condition to be attained for some value of ζ, further vertical motion of the $P-\omega^2$ curve cannot take place; otherwise, if it were to lose contact with the axis altogether, two of the roots of the equation $P=0$ would become complex—an impossibility for the case of a

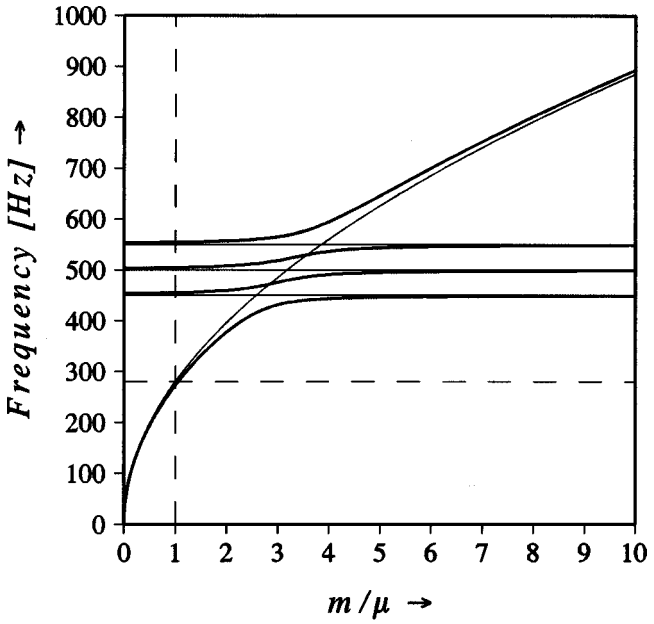


FIG. 2. Hypothetical behavior of the frequencies of three ''wood modes'' and one ''air mode'' as the number of open holes is varied. The abscissa is the ratio of an effective wood mass to the effective air mass, which is linear in the number of open holes.

stable conservative system, whose frequencies must be real. Thus it is necessary, in order to have two eigenvalues coincide, for P to reach an extremum with respect to ζ and ω^2 simultaneously. Such a situation is very unlikely in the first place, and almost certain to be destroyed by any small perturbation of the parameters of the system.[12]

Returning now to Le Gruyère: since Shaw's calculations indicate that the frequency of the $A0$ mode, normally below 300 Hz, can rise beyond 600 Hz as the corks are opened, it is clear that in the process some wood modes will have to be crossed, bringing the phenomenon of mode veering squarely into the picture.

To illustrate the prototypical behavior, we started with Eqs. (3) and (4) assuming (for the sake of argument) that there are three wood modes and one air mode (the $A0$), and solved them (numerically) for the four normal frequencies as the value of μ is changed. For this purpose, we arbitrarily chose the three wood modes to have the same mass coefficient $m_1=m_2=m_3\equiv m$, and stiffness coefficients k_i such as to make their uncoupled frequencies assume the round values of 450, 500, and 550 Hz, respectively. As for the variable mass coefficient for the air, μ, it was set equal to m when the $A0$ frequency is 280 Hz. These assumptions are not meant to be particularly exact; on the other hand, we believe that they are typical within an order of magnitude.

The results are plotted in Fig. 2. The light lines give the results for the case where all three β_i are zero, so no coupling occurs; whereas the heavy lines assume that $\beta_1=\beta_2=\beta_3=0.4$. The abscissa, nominally m/μ, is [by Eq. (7)] proportional to the number of open rib holes; making it go down to zero then corresponds to imagining the *f*-holes, too, as capable of being plugged.

Looking at the light (uncoupled) curves first, we immediately note the three wood modes as horizontal lines. Their radian frequencies are $(k_i/m)^{1/2}$, independent of the number of rib holes that are open. By contrast, what we call the $A0$ frequency varies as the square root of the number of corks removed.

When air-wood coupling is introduced by giving nonzero values to the βs, we obtain the heavy curves of Fig. 2. In this case, the following features [which can, of course, also be obtained by studying Eqs. (3) and (4) directly] are noteworthy:

(a) As expected, no modes ever cross; instead they shift so as to replace one another as the number of open holes increases.
(b) When all holes are completely closed, the ''air mode'' frequency still goes to zero (cf. Sec. A 2 of the Appendix). The ''wood mode'' frequencies, on the other hand, go to values somewhat *larger* than $(k_i/m)^{1/2}$, because the enclosed air now adds to the stiffness of the wood; this increase of frequency is proportional to β^2.
(c) At the point where, in the absence of coupling, the modes would cross, their splitting is largest, being, in fact, proportional to the first power of β.

In Fig. 2, the frequencies of the three wood modes do approach their ''unperturbed'' values (that is, the ones indicated by the light lines) as m/μ becomes large. This is to be expected because, as we already mentioned in Sec. II A, the limit $\mu\rightarrow 0$ is the one in which the wood motion becomes dynamically unaffected by the presence of the air. By contrast, as m/μ approaches zero, the air becomes essentially immobilized, so that (as we have already said) its full bulk modulus contributes to the restoring force of all the modes. One should note, however, that the patterns of wood motion in the resulting normal modes are not, in general, the same as in the ''pure wood modes,'' since the added air stiffness will almost surely mix them.

In the experimental section that follows, we shall use Fig. 2 as a qualitative guide to what we expect to find.

III. EXPERIMENT AND RESULTS

A. General setup

All experiments were performed in the same semianechoic chamber that we have used before.[8] The frame that held the violin was suspended from three light chains in such a way that the instrument was horizontal and near the center of the chamber, supported by a narrow padded ledge at what is normally the chin end, and by a rubber band loop at the neck. (This latter feature is different from the padded but rigid notch in which, in earlier experiments, we used to support the neck, but which we have since found capable of perturbing the violin's individual vibration modes greatly.) The strings were tuned to normal tension and then damped with plastic foam.

In all cases, excitation was by an acoustic signal that originated in one of two loudspeakers, one mounted above the violin facing down, the other mounted below it facing up. The speakers (JBL model 4408) were driven by a computer-generated Schroeder-type quadratic-phase chirp signal[13] con-

taining all frequencies from 120 to 3000 Hz in steps of 1 (more precisely, 1.000 58 Hz) at a sampling rate of 8.196 72 kHz. The data that we actually used in this report extended from 200 to 850 Hz, the range which more or less corresponds to the monopole regime.[8] The computer used for both data acquisition and analysis had an AMD K-6 processor operating at 200 MHz, with a Data Translation DT2821 interface board providing synchronized D/A and A/D conversion with 12-bit precision.

The excitation signals were calibrated by removing the violin and positioning a pair of Primo Model EM-71 microphones in its stead. The microphones were separated vertically by 7.62 cm, the location of their midpoint coinciding with the midpoint between the two bridge feet when the violin was inserted.

The data obtained consisted of the signal received by one of two kinds of sensor pairs, using one of two types of excitation, and one of two ways of combining the signals of each sensor pair. The resulting eight possibilities will now be discussed in detail.

B. Two types of excitation

The calibration of the sound excitation signals by two separate microphones allowed us to calculate, at each frequency, what linear combination of the two would produce an acoustic field which, at the midpoint between the microphones, would have unit pressure amplitude and zero pressure gradient in the vertical direction; or alternatively, one that would have unit (upward) vertical gradient and zero pressure amplitude. We refer to the two as, respectively, a *monopole field* and a *dipole field*, and classify the data accordingly. In fact, of course, we did not attempt to produce those linear combinations of the speaker signals themselves; rather, we relied on the linearity of the system to produce the same result by forming the corresponding linear combinations of the (digitally) recorded sensor output signals.

The nomenclature of ''monopole'' and ''dipole'' is obviously oversimplified for this case, since no attempt was made to determine or adjust gradients in other than the vertical direction. It nonetheless remains useful as a heuristic indication of the two types of fields.

C. Two kinds of sensor pairs

Data were obtained either with a phonograph pickup, with the stylus resting on the bridge halfway between the D and A strings, or with a pair of microphones located inside the violin cavity.

The phonograph pickup was a Stanton D17EE stereo model, mounted on a tone arm which allowed the pickup the usual vertical freedom with constant force but prevented any horizontal motion. Its two channels produced conventional ''RIGHT'' and ''LEFT'' signals, defined by one of the two possible motions at 45° with the vertical. Signals from the pickup were transmitted via short shielded cable to a preamplifier which provided a fixed gain to each channel of approximately 1000, from where longer shielded cables carried the amplified signals to the computer interface board.

The microphones (not to be confused with the ones used for calibrating the excitation sound, Sec. III A) were Knowles Model EK-3033, inserted into the violin cavity through the *f*-holes at the ends of fine shielded cables. They end up being located at the two extreme inside positions: one at the neck end of the violin, the other at the tailpiece button end. As far as we could tell, neither the wires going through the *f*-holes, nor the microphones themselves touching the interior of the violin, produced any appreciable perturbation. The microphone signals were amplified by a preamplifier that had stepped gains from 1 to 2000 before going on to the computer; we found a gain of 200 to be generally satisfactory.

D. Two ways of combining sensor signals

In analyzing the data, the computer was programmed not to deal with the two channels separately but rather with the sum and difference of the signals supplied by those channels. In the case of the output of the phonograph pickup, this effectively converted the two 45°-motions into a horizontal and a vertical direction, which are generally easier to interpret intuitively.

In the case of the data from the internal microphones, taking the sum and difference was especially helpful in that the resulting signals indicate the strength of, respectively, a symmetric and an antisymmetric air vibration. This, as will be seen below, immediately distinguishes the $A0$ mode, which is almost exactly symmetric, from the $A1$, which is antisymmetric.

E. Measurement procedure and display

As stated in Sec. III A, data were obtained at all frequencies from 200 to 850 Hz in steps of (approximately) 1 Hz. At each frequency, eight complex quantities were measured, corresponding to the eight possible permutations of measurement parameters enumerated in Secs. III B, III C, and III D. This procedure was repeated for each of 15 different states of cork removal, beginning with no corks removed, and ending with 28 removed from the lower bout and 14 removed from the upper bout for a total of 42 open holes. The holes were opened in groups of three, always two from the lower bout (one on each side) and one from the upper bout (alternating sides), so as to maintain consistency with Shaw's calculations. The corks were removed beginning at the two ends of the violin and working toward the middle. (The rib holes in the *C*-bouts were never opened, for the reasons to be explained in Sec. IV B.)

For a particular choice of measurement parameters and state of cork removal, the data would normally be displayed as a pair of line graphs giving the amplitude and phase of the corresponding complex number as a function of frequency. An example is shown in Fig. 3, corresponding to the vertical motion of the bridge under monopole forcing when no holes are open. Our problem, however, is to exhibit the behavior of a pair of curves such as those of Fig. 3 when the corks are progressively removed in 15 steps, and a simple juxtaposition of 15 curve pairs does not provide any clear visual impression.

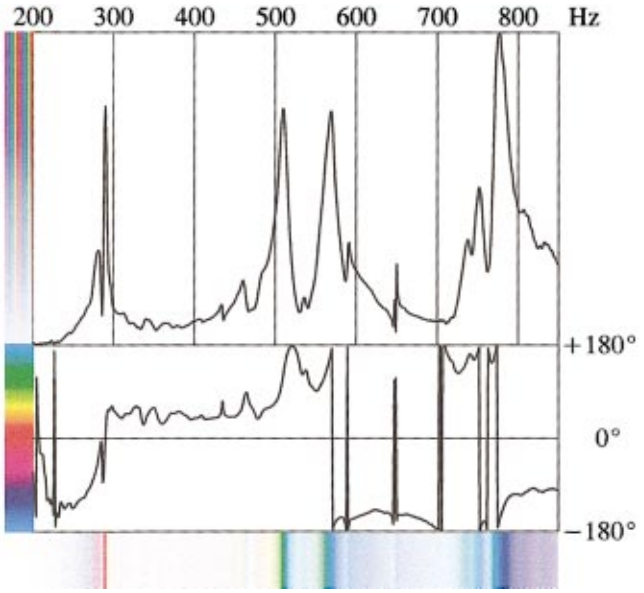


FIG. 3. Amplitude and phase of the vertical response at the midpoint of the bridge, as a function of frequency, when driven with a monopole field and no holes are open. The color conversion code (amplitude=saturation, phase=hue) is shown along the vertical axis, and the resulting colored band for the same data runs at the bottom of the diagram. It is identical to column 0 in the ''WOOD/VER/MONO'' section of Fig. 5.

Accordingly, we chose a different method of display (not, to our knowledge, previously used) which relies on color and works as follows. First of all, the *hue*, by its nature a cyclic variable following the repetitive continuous sequence

red→yellow→green→cyan→blue→magenta→red→…,

is made to represent the sequence of phases

$$0 \rightarrow 60 \rightarrow 120 \rightarrow \pm 180 \rightarrow -120 \rightarrow -60 \rightarrow 0° \ldots .$$

Second, the *color saturation*, which runs from 0% to 100%, is made to represent the amplitude as a percentage of the maximum amplitude attained in the particular graph.

These correspondences are indicated along the left edge of Fig. 3, and the resulting representation of the same data by a one-dimensional band of color is shown along the bottom edge. It is clear that, while this band lacks the visual precision of the line graphs, it does very clearly indicate all the significant features. Indeed, it is gratifying to note how the occasional sudden jumps of phase between +180 and −180°, which are in any case fictitious, disappear in this representation. On the other hand, a resonance always calls attention to itself by having both a relatively high amplitude (= high saturation) and a rapidly changing phase (= rapidly changing hue).

Figures 4 and 5 show the accumulated data, the first from the stereo pickup at the center of the bridge, the second from the two microphones inside the violin. Within each of these the data are grouped in four subsets, depending on whether the driving signal is of the monopole or dipole type (Sec. III B) and on whether the symmetric or antisymmetric combination of the two signal channels is used (Sec. III D). (Because of the reciprocity principle, the data from the stereo pickup, Fig. 4, also represent the radiativity of the violin.[3]) These subsets, labeled with self-explanatory abbreviations, are each plotted in a two-dimensional space with frequency running vertically from 200 to 850 Hz and the number of open rib holes horizontally from 0 to 42 in steps of 3. Thus the color band at the bottom edge of Fig. 3 is identical to the leftmost vertical band in the set labeled ''WOOD/VERT/MONO'' in Fig. 4. It should be noted, however, that a constant saturation scale is used within each subset; in other words, 100% saturation corresponds to the maximum amplitude in the set of 15 vertical bands, and not within each band separately.

IV. DISCUSSION

A. The air modes: *A*1

In order to tie in with the work of Shaw,[6] we wish first to identify the $A0$ and $A1$ modes; as expected, they dominate the monopole-driven microphone signals in their symmetric and antisymmetric combinations, respectively (Fig. 5). For comparison, Fig. 6 shows Shaw's prediction for the variation of these two mode frequencies with the number of open holes, plotted to cover the same range as our data. This graph is based on Fig. 12 of Ref. 6, except that we have scaled the abscissa to account for the correct average dimensions of the holes (Sec. II A).

Looking at the experimental $A1$ frequency first, we note that its qualitative behavior coincides well with the theoretical prediction. Quantitatively, however, there are two differences:

(a) The zero-open-hole observed frequency is about 485 Hz rather than Shaw's 517 Hz, a difference of 32 Hz. This is as it should be, however, because Shaw made his calculations for the case of rigid plates, whereas our plates are, of course, free. Shaw's value corresponds well, in fact, to Hutchins's measurements with the instrument covered with sand and weighted down with lead, while leaving access to the *f*-holes open; whereas ours is in good agreement with her measurement when the instrument is free[6] (but note the problem pointed out in Sec. IV E).

(b) There is a much larger discrepancy, however, at the maximum (42) number of open holes: Shaw predicts about 790 Hz for the $A1$, whereas we measure about 650 Hz, smaller by 140 Hz. Although no direct data are available, it is hard to imagine why the effect of wood nonrigidity would here be more than quadrupled compared to the situation with all rib holes closed.

B. Variation of mode pressure with position

There is another factor, however, which Shaw's paper noted but did not pursue, and that is the variation of pressure amplitude along the upper and lower bouts when the air is vibrating in the $A1$ mode. In particular, Shaw's lumped-circuit model takes the upper and lower bouts as separate volumes connected by an inertive channel, with additional inertive parallel channels supplied by the *f*-holes and the open rib holes (his Fig. 8). According to this picture, it makes no difference *where* a rib hole is opened in one of

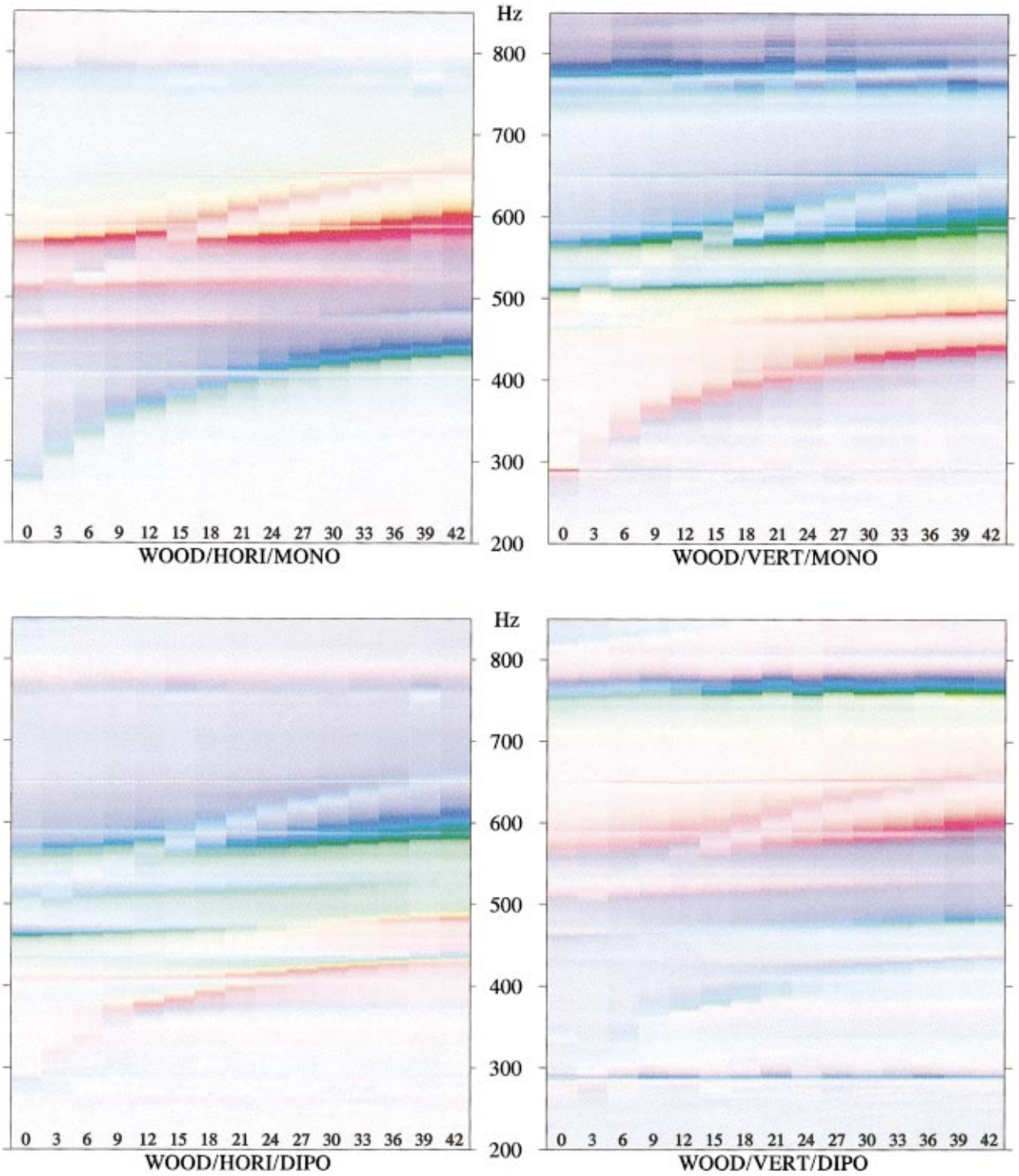


FIG. 4. Wood response: measured motion of the midpoint of the bridge of Le Gruyère in two directions (horizontal and vertical) and under the action of two types of acoustic forcing fields (''monopole'' and ''dipole''). Within each of the four subgraphs, the abscissa is the number of open rib holes and the ordinate is the frequency. Amplitude and phase are indicated by the scheme of Fig. 3.

those bouts. One might expect such an approach to work fairly well for the $A0$ mode, but we know that the $A1$ has a nodal surface somewhere in the middle, and the pressure can hardly drop from a constant amplitude to zero completely suddenly.[9]

Another way of describing the same effect is to note that for the $A0$ mode the flow of air within the wooden shell is very small except right near the *f*-holes, but that this is not so for the $A1$. Therefore, for the latter case, the capacitances representing the volumes of the bouts ought to be subdivided, with some inertance inserted in series between the subdivisions.

In order to investigate this question experimentally, we measured the shift in the frequencies of the $A0$ and $A1$ modes when the rib holes are opened just one at a time, as a function of which hole is chosen. Opening such a hole introduces, in lowest order, an additional amount of kinetic energy (due to the air flowing through the opened hole) without

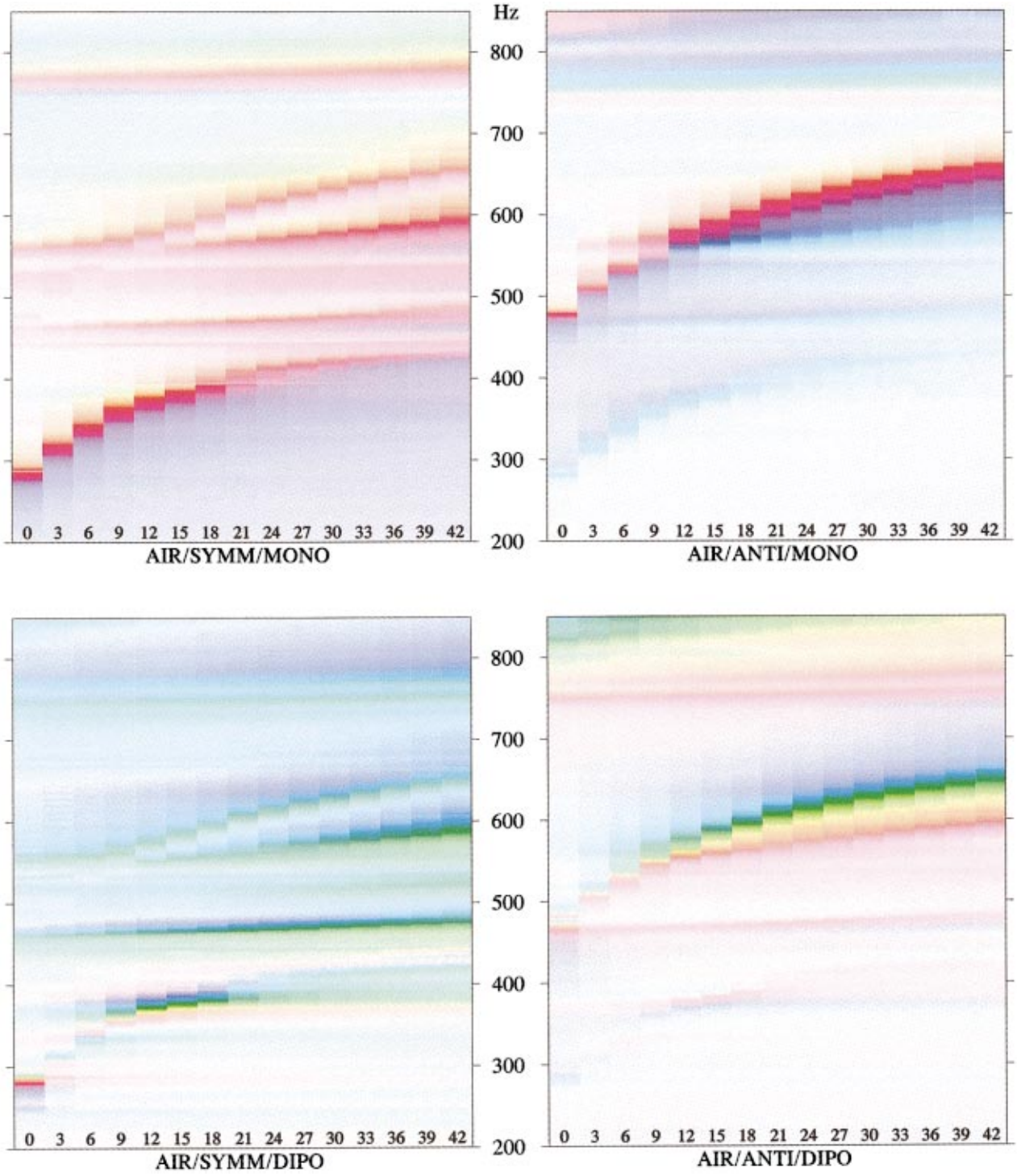


FIG. 5. Air response: as in Fig. 4, but the signals come from two small microphones at opposite ends of the violin air cavity. The sum of the two (separately calibrated) responses is marked ''SYMM,'' and the difference is ''ANTI.''

affecting the potential energy, thus lowering the frequency; this additional kinetic energy is, clearly, proportional to the square of the pressure at the position of the hole when the (normalized) mode is excited. Therefore, measuring the frequency shift as a function of which hole is opened ought to produce a ''map'' of the square of pressure amplitude as a function of position.

One does, however, need to make a further correction if the individual holes are not of identical size. Specifically, when a hole is opened, the resulting volume velocity u is given in terms of the mode pressure p by

$$u = p/L_h\,, \tag{10}$$

L_h being the inertance of the hole;[7] and the associated kinetic energy change is

$$\Delta T = \tfrac{1}{2} L_h u^2 = \tfrac{1}{2} p^2 / L_h\,. \tag{11}$$

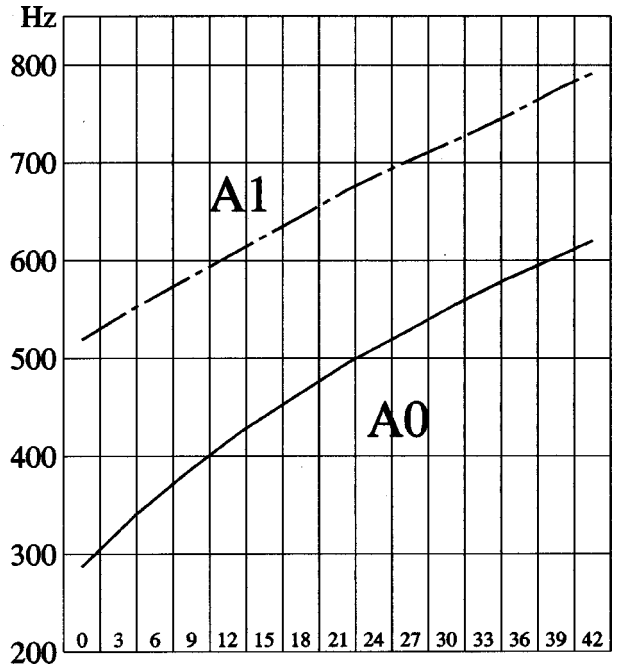


FIG. 6. Variation of the frequencies of the $A0$ and $A1$ modes with the number of open rib holes as calculated by Shaw (Ref. 6, Fig. 12b), corrected for actual hole dimensions, and drawn in the same manner and to the same scale as the experimental results of Figs. 4 and 5.

Since Δf, the shift in frequency, is proportional to ΔT, it follows that

$$p^2 \propto L_h \Delta f; \tag{12}$$

and using the formula for the inertance of a hole of radius r in a wall of thickness t quoted by Shaw in his Eq. (49), we obtain

$$p^2 \propto [(t+1.7r)/r^2]\Delta f. \tag{13}$$

According to our measurements of the holes of Le Gruyère, t is quite variable, apparently because the drilling operation often left an inward ''chimney'' surrounding the hole; but since the formula (13) is relatively insensitive to t, we assigned it a constant value of 1.1 mm based on an approximate average. As for the diameter $2r$, it varies (as already mentioned in Sec. II A) between 5.81 and 8.32 mm, and we used the appropriate measured value for each individual hole.

Figure 7 shows our results for p^2, in arbitrary units, according to Eq. (13). The top graph is for the $A0$ mode, the bottom one for the $A1$; in each case, the holes are numbered beginning closest to the tailpiece button and ending at the neck. Crosses indicate the treble side of the violin and circles the bass side. The values of Δf were determined by fitting a parabola to the three highest points of the AIR/SYMM/MONO graph (for the $A0$) or the AIR/ANTI/MONO (for the $A1$).

Although we have not attempted to build a definite theoretical model for this phenomenon, it would appear that the drop in $A1$ mode pressure between the extreme positions (that is, button and neck) and the beginning of the C-bouts easily accounts for the fact that the removal of corks near the C-bouts has a decidedly weaker effect on the $A1$ frequency,

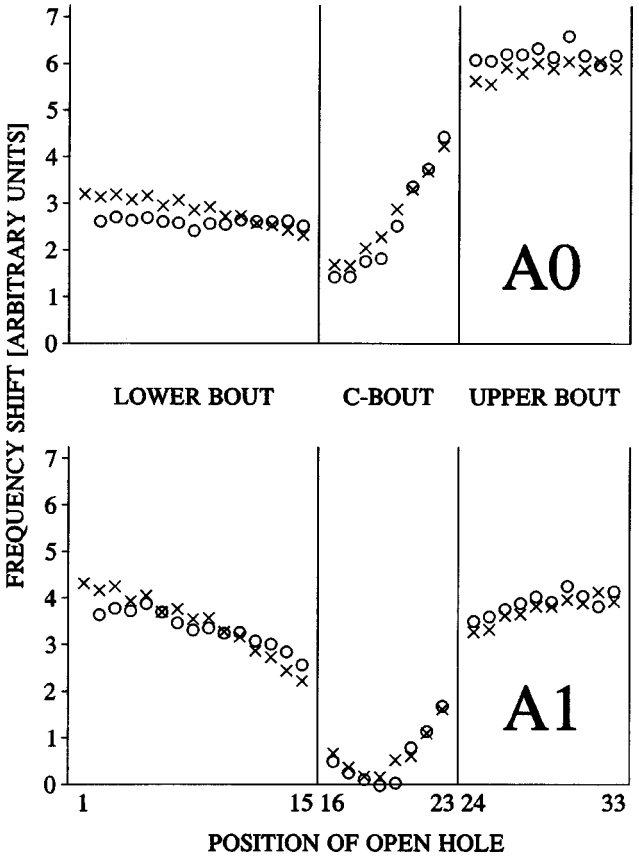


FIG. 7. Shift in frequency of the $A0$ mode (top) and the $A1$ mode (bottom) when holes are opened one at a time, as a function of which hole is chosen. The frequency shift is corrected for variation in hole size, and is hence plotted in arbitrary units; the result is proportional to the squared pressure amplitude of the normalized mode at the position of the hole. The holes are arbitrarily numbered beginning at the button and ending at the fingerboard, the side of the violin being indicated by crosses (treble side) or circles (bass side).

thus explaining why the frequency rise of that mode is not as large as Shaw calculated. This is also the reason why, in our main data sequences, we chose not to go beyond 42 open holes, or 14 from the upper bout and 28 from the lower: beyond this point, we would be removing corks from the C-bouts themselves, a situation which Shaw's work does not attempt to cover.

C. The air modes: *A0*

By contrast with the $A1$, the graph for $A0$, also shown in Fig. 7, indicates a more or less constant mode pressure on each side of the C-bout. The frequency shifts do differ quite a bit for the upper and lower bouts; specifically, the former is higher than the latter by a factor of about 2.1, corresponding to a pressure ratio of about 1.45. This is somewhat larger than 1.30, which is the ''adjusted'' value calculated by Shaw [his Eq. (46)], though coming quite close to his ''unadjusted'' value of 1.43 [his Eq. (42)]; either way, we consider the agreement satisfactory. Thus we would expect that, since the pressures are more or less constant for each bout sepa-

rately, and since we always remove two corks from the lower bout and one cork from the upper bout, the agreement with Shaw's calculation ought to be good.

Yet this is not so: comparison of Fig. 6 with Figs. 4 and 5 shows, in the latter (experimental) case, an $A0$ frequency which begins in the right place (about 280 Hz) and at first rises more or less at the expected rate, but then flattens off radically and reaches the right edge of the diagram at about 430 Hz rather than the predicted 600 Hz or so of Fig. 6; this mode is seen most clearly in ''WOOD/HORI/MONO'' and ''AIR/SYMM/MONO.'' At the same time, there is another mode which begins rather horizontally at about 568 Hz and bends in the second half of its range to reach just about the 600 Hz predicted by Shaw. This mode appears even more strongly in the monopole radiativity for horizontal bow motion (WOOD/HORI/MONO), revealing itself as an important contributor to the violin's sound radiation.

We suggest that what is happening here is that the originally unperturbed $A0$ mode, as calculated by Shaw, is strongly modified by its interactions with two wood modes. In the absence of such interaction, one of these wood modes (probably the ''$B1-$'') would be at about 464 Hz, the other (probably the ''$B1+$'') at about 568 Hz.[14] With the interaction present, the rising $A0$ first runs into the $B1-$, which causes the $A0$ to flatten out and send what was previously a wood mode upward in its stead; the same phenomenon is then repeated as the new mode approaches the $B1+$. In essence, as $A0$ flattens out, it takes over the ''wood mode'' characteristics of $B1-$, while the former $B1-$, as it rises, receives the ''air mode'' characteristics of $A0$; a similar interchange then occurs as this new rising mode encounters the $B1+$.

Figure 8 shows a computed prediction based on such a model. The thin lines represent the two wood modes at the frequencies mentioned in the last paragraph, plus the $A0$ which, in this simplified picture, rises inversely with the square root of μ. The units of this ''effective air mass'' are defined by setting $\mu=1$ at 280 Hz, the known value of the $A0$ when all rib holes are closed. Although the graph shows $1/\mu$ varying from 0 to 7, this whole range is not, of course, experimentally accessible: its lower limit is determined by the existence of open *f*-holes, the upper one by the number of rib holes that have been drilled. (This range is indicated in the graph as the unshaded area, whose vertical extent is drawn to match the frequency span covered by the data graphs of Figs. 4 and 5.) By using Shaw's estimate for the inertance of the *f*-holes and rib holes, we obtain a ratio of maximum to minimum $1/\mu$ very close to 4, the value used in this graph.

On the same graph, the heavy lines indicate the numerically computed frequencies of the three modes when they are coupled according to Eqs. (3) and (4), using the parameters $m_1=0.2$, $m_2=0.3$, $\beta_1=0.1$, $\beta_2=0.1$. The resulting pattern is rather similar to the experimental one. No claim is here made that these particular values are any kind of ''best fit,'' especially since the curve representing the uncoupled $A0$ is based on the simplified square-root model which ignores the series inertance of the channel between the two bouts. Figure 8 merely illustrates how easily the data can be fitted with what appear to be entirely reasonable parameters.

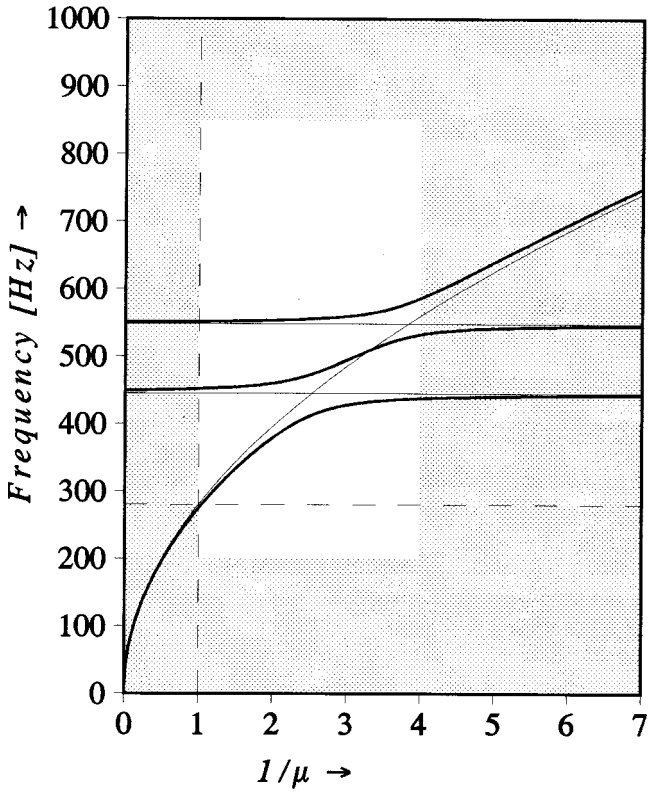


FIG. 8. Calculated frequencies of three coupled modes (one ''air mode'' and two ''wood modes'') assuming the parameters of Sec. IV C, as a function of the reciprocal of the effective mass of the air mode (which is proportional to the number of open rib holes of the violin). The unshaded rectangle corresponds to the range covered by the graphs of Figs. 4 and 5.

D. ''Polarity of resonance''

In a paper of this title a third of a century ago, Schelleng[15] first discussed the importance of the *sign* (or ''polarity'') of the modal monopole radiativity in determining the radiation behavior of a violin when more than one mode is active, and pointed out that the polarities of the ''main air resonance'' and ''main wood resonance'' must be opposite to each other (as would also be expected from the sound hole sum rule[3]). It is worth noting how our data make this fact immediately obvious. In particular, the WOOD/HORI/MONO plot of Fig. 4 shows the horizontal monopole radiativity of the violin, that is, the monopole radiativity in the bowing direction. Clearly, the ''main air resonance'' (which begins around 280 Hz) is generally blue in color, which means that its phase is around 180°; whereas the ''main wood'' cluster around 570 Hz is generally red, which means a phase near zero. Although not central to the present work, this observation does illustrate the value of being able to keep both amplitude and phase information in our data displays.

E. Coupling of the *A*1 mode

As already pointed out in Sec. IV A, the $A1$ mode is most strikingly visible—as one might expect—in the graph

of Fig. 5 labeled AIR/ANTI/MONO, beginning at about 485 Hz and ending about 655 Hz. The graph just to the left of it, labeled AIR/SYMM/MONO, shows no trace of its beginning, in agreement with Shaw's prediction that the pressure amplitude of this mode in the upper and lower bouts is the same. This situation changes, however, when enough corks are removed to make the $A1$ ''cross'' the wood mode $B1+$ which begins at about 568 Hz and which, as we already said, interacts strongly with the $A0$. Once that mode is crossed, the emerging $A1$ apparently acquires unequal pressure amplitudes in the two bouts, so that it becomes visible in the SYMMetric graph as well as the ANTIsymmetric one.

The same mode crossing is also visible in all four graphs of Fig. 4; this time it is the wood mode which is prominent, and the $A1$ which is subdued. In all cases, it is clear that the coupling of the $A1$ to the wood mode is weak, explaining why the $A1$ does not generally contribute strongly to the sound radiation of the violin.

There is, however, a bit of a mystery here, since according to Hutchins the effect of air-wood coupling (as measured by comparison with a situation in which the wood of the violin is immobilized) is to lower the frequency of the $A1$ by 40 Hz (from 528 to 488 Hz) while actually *raising* the frequency of the $A0$ by 3 Hz (from 287 to 290 Hz).[16] Neither one of these results seems consistent with our measurement of coupling. For the $A1$, the graphs suggest a maximum splitting where the $A1$ crosses the wood mode of perhaps 15 Hz; since this quantity is of first order in the interaction parameters, and the shift with all holes closed is of second order, one would expect the $A1$ to change by no more than perhaps 1 Hz! Conversely, for the $A0$ mode the graphs suggest a first-order splitting (at the point where the $A0$ and the wood mode are most strongly veering away from each other) of perhaps 50 Hz, leading us to expect a second-order shift (when all holes are closed) of the order of 10 Hz or more. Furthermore, this last shift, which is the shift of the $A0$ frequency due to the nonrigidity of the wood, ought certainly to be in a downward direction, since all the wood modes with which it interacts are above it. We know of no explanation of this mystery, except perhaps that Hutchins' experiment of immobilizing the wood is surely a very difficult one, and some technical problems may have crept in.

F. Other modes

Careful examination of the graph AIR/SYMM/MONO (Fig. 5) reveals a rather sharp mode at about 290 Hz visible when all holes are closed but disappearing immediately thereafter. This is undoubtedly the $B0$ mode, in which the whole instrument vibrates rather like a flexing plate, the scroll and button moving in one direction while the middle part (including the fingerboard, and especially its cantilevered section) moves the other way.[17] That it disappears from Fig. 5 as suddenly as it does indicates that it was being excited in the first place via its (rather weak) coupling to the $A0$ only as long as the two were close in frequency.

To excite it directly, one would do better looking at the wood motion itself. The bridge motion associated with this mode is pretty much vertical, and it is more sensitive to dipole fields than to monopole fields, not being associated with any great volume change. Indeed, the WOOD/VERT/DIPO graph of Fig. 4 shows it as a clear horizontal straight line moving horizontally across the diagram. It is also visible, in a weaker form, in the WOOD/VERT/MONO graph, showing that its volume change is not exactly zero. In fact, it is here most distinct at the left edge of the diagram, where the $A0$ is at a frequency such as to be strongly coupled to it.

A number of additional very sharp wood modes cross the diagrams of Figs. 4 and 5 horizontally, without appreciable coupling to anything else. We see such modes, for example, at about 652, 586, and 410 Hz. Presumably they have little effect on the sound of the violin.

V. CONCLUSIONS

The theoretical method of ''dynamic modes'' introduced in this paper, which is generally applicable to the coupled vibration of fluids and their enclosures, has been successfully applied to the Swiss-cheese violin. In the process, we were able to reinterpret the calculated results of Shaw for the case where the wooden shell of the violin is flexible. Measurements of various response functions of the system as the number of open rib holes is varied, displayed in terms of a novel color code that takes into account both amplitude and phase of the measured quantity, clearly exhibit the ''mode veering'' which occurs when the frequencies of strongly coupled modes attempt to cross. This experimental behavior can be easily matched with reasonable system parameters.

ACKNOWLEDGMENTS

Our most heartfelt thanks go to the following colleagues who, by graciously taking the time to read through a draft of the manuscript and giving us their thoughtful comments, have improved the paper immeasurably: George Bissinger, Xavier Boutillon, Joseph Curtin, Oliver Rodgers, Edgar Shaw, and James Woodhouse. Obviously, we acknowledge any remaining blunders as our own. We are grateful also to Sharon Que for allowing us to use her masterful drawing of a violin which underlies Fig. A1. Last but not least, we express our appreciation to the National Science Foundation, which partially supported this research.

APPENDIX: MONOPOLE REGIME

1. Description of the regime

As is well known, sound emission from a space-limited source at low frequencies has the form of monopole radiation; that is, (a) it is isotropic, and (b) it depends only on the *monopole moment* of the source, that is, on the amplitude of its overall volume fluctuation. The mathematical reason for this behavior is, of course, that the multipole expansion of the radiation field (which converges for all frequencies) is a power series in the quantity kR, where k is the magnitude of the propagation vector at that frequency, and R is the radius of a sphere large enough to encompass the whole source. Thus at sufficiently low frequencies, where kR is a small quantity, the first (monopole) term necessarily dominates. An exception to this behavior occurs only if the monopole mo-

ment happens to vanish identically, in which case the next nonzero term will dominate—but the radiation will naturally be much smaller.

For a spherical source of radius a, the monopole limit—that is, the approximate frequency below which monopole behavior dominates—is usually defined by the condition $ka=1$, which means that the circumference of the sphere is equal to λ, or half the circumference to $\lambda/2$. The last formulation is, perhaps, the most intuitive, if one thinks of the essential physics of the monopole regime in terms of what is often referred to as the ''short-circuiting'' of higher moments. Specifically, imagine a sphere whose surface is displaced in non-monopole fashion, that is, in such a way that any outward displacement at one point is accompanied by an equal inward displacement somewhere else. If the circumferential distance between those two points is less than is required for the wave to accumulate an appreciable phase shift, the two partial sources will cancel each other strongly. Now the farthest that any two points on a sphere can be from each other is half the circumference, and the distance causing a full phase reversal is half a wavelength. Comparing those two quantities, the monopole condition immediately follows.

Such a way of thinking about the problem is especially useful when we try to extend our discussion to a source shape which is not spherical, such as a violin. The relevant length to be compared with half a wavelength is then constructed as follows: First, we determine the shortest surface distance $s(P,P')$ connecting any given pair of surface points P and P'; second, we find the maximum value which s takes as P and P' are varied. If this maximum distance is smaller than half a wavelength, it is impossible to construct a non-monopole displacement of the surface which could radiate without being short-circuited. We emphasize again that this is not a black-or-white condition, but only an indication of the frequency at which the transition to monopole radiation is strong. (Strictly speaking, the choice of P and P' should be weighted by the degree to which those points actually vibrate.)

The behavior of a violin is a good illustration: naively asking for an ''average radius'' of a violin might result in an estimate of perhaps 12 cm (Fig. A1). If the violin were a sphere of that radius, the frequency that would make half the circumference equal to half a wavelength would be about 455 Hz. Empirically, however, the monopole limit for a violin is about 850 Hz,[8] corresponding to an ''effective radius'' of about 7 cm. In terms of the argument of the preceding paragraph, the reason for this discrepancy is clear: due to the existence of the *C*-bouts, 7 cm is not an unreasonable estimate for $1/\pi$ times the shortest distance between the center of the top and the center of the back.

2. Number of violin air modes in monopole regime

Since the factors that determine the frequencies of a violin's wood modes have nothing to do with wave propagation in air, the number of such modes below the monopole limit is not *a priori* determined; but this is not the case for the air cavity modes. Consider, first, a completely enclosed spherical cavity. Of the various possible standing wave modes that can exist in it, the lowest is triply degenerate and has a frequency given by $ka=2.081575$, considerably above unity.[18] Hence we conclude that, for a closed spherical cavity, *no standing wave mode exists in the monopole regime*.

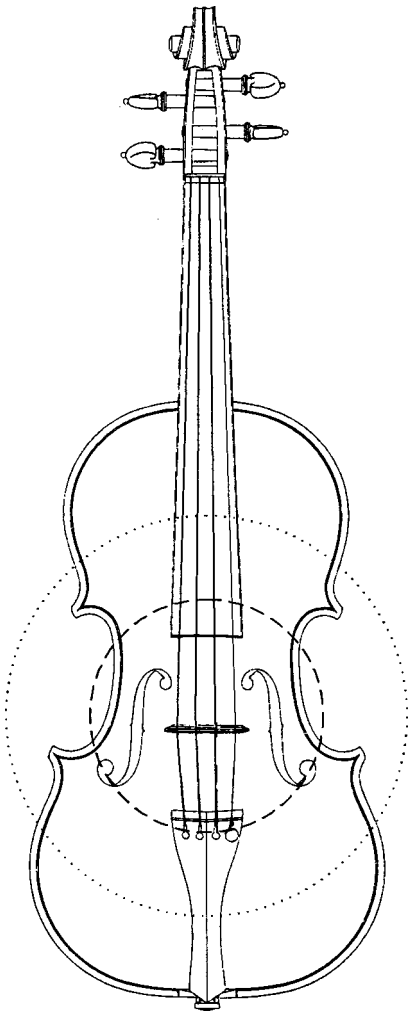


FIG. A1. Equivalent monopole sphere of a violin. The dotted circle indicates a 12-cm radius, which one might expect visually. The dashed circle is the empirically measured 7 cm.

Of course the situation is different if the wooden shell has one or more ''sound holes,'' as is true for string instruments. In this case a new mode, the ''Helmholtz mode'' or ''$A0$,'' appears; this is the mode in which there is an oscillating flow in and out of the sound hole(s), with the air in the hole(s) supplying the inertia and the air in the rest of the shell acting as a spring. For a spherical shell, the frequency of the $A0$ is, in order of magnitude, smaller than the monopole limit by $(R/a)^{1/2}$, where R is the radius of the hole (assumed circular). Hence, for a reasonable hole size, we conclude that there will be one and only one mode, the $A0$, within the monopole regime. (Strictly speaking, the $A0$ does not appear *ex nihilo* when the hole is opened. Rather, it arises from the zero-frequency mode for a hermetically closed shell, which simply expresses the fact that if the pressure is raised above ''normal'' in such a shell, the overpressure will not change with time.)

Passing now to the case of the nonspherical violin shell, two changes have to be introduced. First of all, as we already saw, the presence of the *C*-bouts *raises* the monopole limit,

opening the possibility that air modes other than the $A0$ might now fall within it. Second, the same *C*-bouts also *lower* the frequency of the so-called $A1$ (which is the air mode next above the $A0$ in frequency), as follows. The $A1$ derives from one of the triply degenerate sphere modes, and can be visualized as a ''sloshing'' of air back and forth between the upper and lower bouts. If the violin had the shape of a cylindrical pipe, the $A1$ frequency would, of course, be given by $kL=\pi$, L being the length of the pipe; taking the (internal) length as about 30 cm would then predict it to be about 575 Hz. But the constriction introduced by the *C*-bouts also introduces an additional amount of inertia, which lowers the $A1$ frequency by about 20%, typically to 475 Hz. (One can appeal here to Rayleigh's fundamental argument[19] that constricting a pipe in a region of kinetic, rather than potential, energy always lowers the frequency of a mode; alternatively, one can think of the $A1$ as a mode of two Helmholtz resonators attached mouth to mouth, knowing that constricting the mouth of such a resonator always lowers its frequency.)

According to the data of Hutchins,[5] there is a further lowering of the $A1$, due to nonrigidity of the wooden shell, amounting to about 40 Hz. For reasons discussed in Sec. IV E, however, we find this observation problematic.

The net result is that, for a violin, there will be not one but two air modes, the $A0$ and the $A1$, within the monopole regime. Nonetheless, one expects the radiativity of the $A1$ to be small, since the *f*-holes, through which the oscillation communicates with the outside radiation field, are located near a pressure node of this mode; our data confirm this expectation (Sec. IV A). For this reason, the ''quasispherical'' model, in which $A0$ is the only air mode below the monopole limit, often provides a useful approximation.

[1] Throughout this paper we treat our systems as dissipationless. Although losses can be introduced in a rigorous manner, we do not feel that, for our present purposes, the resulting gain would justify the added theoretical complication.

[2] See, for example, F. Fahy, *Sound and Structural Vibration* (Academic, New York, 1985), Sec. 6.4.

[3] G. Weinreich, ''Sound hole sum rule and the dipole moment of the violin,'' J. Acoust. Soc. Am. **77**, 710–718 (1985).

[4] C. M. Hutchins, ''A violin with 65 holes in its ribs, phase II,'' Paper K3-5, Proc. 12th International Congress Acoustics, Toronto, 1986.

[5] C. M. Hutchins, ''A study of the cavity resonances of a violin and their effects on its tone and playing qualities,'' J. Acoust. Soc. Am. **87**, 392–397 (1990).

[6] E. A. G. Shaw, ''Cavity resonance in the violin: Network representation and the effect of damped and undamped rib holes,'' J. Acoust. Soc. Am. **87**, 398–410 (1990).

[7] For a good discussion of inertance (sometimes also called ''acoustic mass'') in such a situation, see A. D. Pierce, *Acoustics: An Introduction to its Physical Principles and Applications* (Acoustical Society of America, New York, 1989), pp. 345–348.

[8] G. Weinreich, ''Directional tone color,'' J. Acoust. Soc. Am. **101**, 2338–2346 (1997).

[9] Shaw did follow up on this question in a letter to Oliver Rodgers dated November 19, 1996 (E. A. G. Shaw, private communication).

[10] O. Rodgers, ''Another look at the Swiss cheese violin,'' Catgut Acoust. Soc. J. **3**(4) (Series II), 17–23 (1997).

[11] N. C. Perkins and C. D. Mote, ''Comments on curve veering in eigenvalue problems,'' J. Sound Vib. **106**, 451–463 (1986).

[12] The situation is different where the eigenvalues can be complex: see G. Weinreich, ''Coupled piano strings,'' J. Acoust. Soc. Am. **62**, 1474–1484 (1977).

[13] M. Schroeder, ''Synthesis of low-peak-factor signals and binary sequences with low autocorrelation,'' IEEE Trans. Inf. Theory **16**, 85 (1970).

[14] We use the very logical nomenclature introduced by George Bissinger: ''Some mechanical and acoustical consequences of the violin soundpost,'' J. Acoust. Soc. Am. **97**, 3154–3164 (1995); ''Ambiguity in $A1$-$B1$ Δ criterion for violin tone and quality,'' Catgut Acoust. Soc. J. **3**(1) (Series II), 13–16 (1996).

[15] J. C. Schelleng, ''On polarity of resonance,'' Catgut Acoust. Soc. Newsletter No. 10, 14–18 (1968); reprinted in *Musical Acoustics part 2*, edited by Carleen Hutchins (Dowden, Hutchinson & Ross, Stroudsburg, PA, 1976), pp. 156–160.

[16] See C. M. Hutchins, ''A study of the cavity resonances of a violin and their effects on its tone and playing qualities,'' J. Acoust. Soc. Am. **87**, 392–397 (1990), Fig. 8(b); E. A. G. Shaw, ''Cavity resonance in the violin: Network representation and the effect of damped and undamped rib holes,'' *ibid.* **87**, 398–410 (1990), Table II.

[17] J. Woodhouse, ''The acoustics of '$A0$-$B0$' mode matching in the violin,'' Acust. Acta Acust. **84**, 947–956 (1998).

[18] See, for example, G. Arfken, *Mathematical Methods for Physicists*, 3rd ed. (Academic, New York, 1985), p. 632. (The reference to a ''Helmholtz resonator'' is, of course, wrong.)

[19] J. W. Strutt (Baron Rayleigh), *The Theory of Sound*, 2nd ed. (Dover, New York, 1945), Vol. 2, Sec. 265, pp. 66–67.

Frequency dependence of ultrasonic backscattering in cancellous bone: Autocorrelation model and experimental results

S. Chaffaï and V. Roberjot
Laboratoire d'Imagerie Paramétrique, UMR CNRS 7623, Université Pierre et Marie Curie, 15 rue de l'Ecole de Médecine, 75006 Paris, France

F. Peyrin
Laboratoire d'Imagerie Paramétrique, UMR CNRS 7623, Université Pierre et Marie Curie, 15 rue de l'Ecole de Médecine, 75006 Paris, France; Creatis, UMR, INSA 502, 69621 Villeurbanne Cedex, France; and ESRF, BP 220, 38043 Grenoble Cedex, France

G. Berger and P. Laugier[a)]
Laboratoire d'Imagerie Paramétrique, UMR CNRS 7623, Université Pierre et Marie Curie, 15 rue de l'Ecole de Médecine, 75006 Paris, France

(Received 14 March 2000; accepted for publication 15 August 2000)

The goal of this study is to model the frequency dependence of the ultrasonic backscatter coefficient in cancellous bone. A twofold theoretical approach has been adopted: the analytical theoretical model of Faran for spherical and cylindrical elastic scatterers, and the scattering model for weakly scattering medium in which the backscatter coefficient is related to the autocorrelation function of the propagating medium. The ultrasonic backscatter coefficient was measured in 19 bone specimens (human calcaneae) in the frequency range of 0.4–1.2 MHz. The autocorrelation function was computed from the three-dimensional (3D) microarchitecture measured using synchrotron radiation microtomography. Good agreement was found between the frequency dependence of the experimental ($f^{3.38\pm0.31}$) and autocorrelation modeled ($f^{3.48\pm0.26}$) backscatter coefficients. The results based on Faran theory (cylindrical Faran model: $f^{2.89\pm0.06}$ and spherical Faran model: $f^{3.91\pm0.04}$) show qualitative agreement with experimental data. The good prediction obtained by modeling the backscatter coefficient using the autocorrelation function of the medium opens interesting prospects for the investigation of the influence of bone microarchitecture on ultrasonic scattering. © *2000 Acoustical Society of America.* [S0001-4966(00)04011-X]

PACS numbers: 43.80.Ev, 43.80.Qf, 43.80.Vj [FD]

I. INTRODUCTION

Osteoporosis is characterized by a decrease of the strength of the bone matrix and an increase in the risk of fracture. This fragility is due to both a loss of bone mass and a deterioration of trabecular bone microarchitecture. Nowadays, x-ray-based densitometry techniques are considered as the gold standard for skeletal status assessment. Only bone mass reduction can be appreciated in current practice using these x-ray absorptiometry techniques. Recent advances of high-resolution x-ray computed tomography (Müller *et al.*, 1996), magnetic resonance microimaging (Gordon *et al.*, 1997; Majumdar *et al.*, 1996), and radiographic images processing (Benhamou *et al.*, 1994) have shown the potential for *in vivo* microarchitecture characterization. However, the alteration of bone microarchitecture cannot be evaluated yet in clinical practice.

Fifteen years ago, the use of quantitative ultrasound (QUS) techniques was suggested as an alternative to x-ray densitometry techniques for several reasons: it is a nonionizing, easy to implement, and less expensive technology (Glüer, 1997). Current available technologies are based on measurements in transmission of the slope of the frequency-dependent attenuation (BUA, dB/MHz) and the speed of sound (SOS, m/s) at peripheral skeletal sites (calcaneus, finger phalanges). The QUS approach in osteoporosis has been clincially validated: the predictive value of QUS for fracture risk at the femoral neck and spine has been documented in several prospective studies (Bauer *et al.*, 1997; Hans *et al.*, 1996; Thomson *et al.*, 1998).

However, the potential of QUS for bone characterization is still not fully exploited. Transmission measurements only partially exploit the information related to the interaction between the elastic wave and bone microarchitecture. Several studies concluded that ultrasonic measurements on human cancellous bones reflected primarily bone mineral density (BMD) and that there was only very limited room for microstructural factors to play a role in current QUS measurements (Laugier *et al.*, 1997a; Nicholson *et al.*, 1998; Hans *et al.*, 1999). In addition, these QUS techniques are limited to parts of the skeleton which are easily accessible for transmission measurements, i.e., only peripheral skeletal sites.

The assessment of bone strength demands increasing knowledge not only about bone mineral density, but also about microscopic bone structure. Ultrasonic tissue characterization using the analysis of backscattered signals has shown the potential to provide the information needed for the assessment of tissue microstructure (Insana *et al.*, 1990;

[a)]Author to whom correspondence should be addressed; electronic mail: laugier@lip.bhdc.jussieu.fr

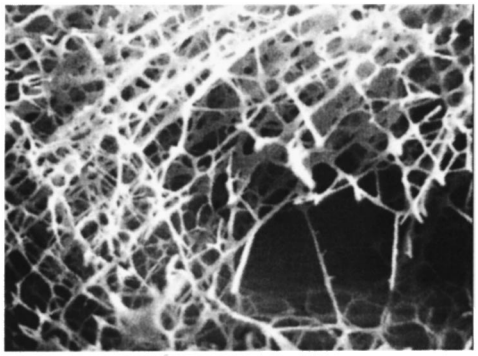


FIG. 1. Illustration of the trabeculae structure of a human calcaneus.

Shung and Thieme, 1993). A backscattering measurement technique exploiting scattering from the internal microarchitecture of bone should give more information about this microarchitecture. In addition, skeletal sites, which are difficult to reach by transmission, such as the femoral neck, could be evaluated by the backscattering technique. Some studies showed the *in vitro* and *in vivo* feasibility of this approach on the calcaneus (Roberjot *et al.*, 1996a; Wear and Garra, 1998), and our group has recently demonstrated that ultrasound backscatterer measurements have adequate ability to discriminate patients with osteoporosis from controls (Laugier *et al.*, 1997b; Roux *et al.*, 1999).

Initial theoretical studies of ultrasonic scattering by trabecular bone were performed by Wear (1999). The approach adopted in this previous study consisted in assuming that elementary scatterers in trabecular bone could be modeled as long cylinders with small diameter compared with wavelength aligned perpendicularly to the ultrasound beam axis. The objective of this paper was to carry this work further and to apply the autocorrelation models to bone microstructure and to compute the backscatter coefficient. Toward this end, the frequency dependence of ultrasonic backscatter from trabecular bone was experimentally investigated and compared with a theoretical model similar to the one developed by Wear (1999) and the new approach based on the autocorrelation function.

II. THEORY

Cancellous bone can be considered as a highly inhomogeneous scattering medium: a fluidlike medium (marrow) containing a solid matrix (mineralized collagen) of interconnected trabecular elements (trabeculae) of mean thickness ranging between 50 and 150 μm (Fig. 1). In this study, the trabecular network, which is supposed to be the main source of ultrasound scattering, is considered as an ensemble of small discrete scattering particles. In addition, the assumption of no multiple scattering was made.

Several theoretical methods exist to model backscattering. A first approach consists of using the analytical model of Faran to describe the scatterers and their interaction with the ultrasonic wave (Faran, 1951). This theory provides an exact solution for the backscatter cross section by a spherical or a cylindrical solid elastic object. The scatterer is assumed to be isotropic, insonified by a plane wave, and embedded in a nonviscous fluid medium. When scattering is observed from a volume containing a random distribution of scatterers, the backscatterer coefficient may be obtained by multiplying the differential backscatter cross section of each scatterer by the number of scatters per unit volume. The expressions of the differential backscatter cross sections for a spherical and a cylindrical scatterer are given, respectively, by (Faran, 1951)

$$\left.\frac{d\sigma_{\text{Sphere}}}{d\Omega}\right|_{\theta\cong 180^\circ} = \left(\frac{1}{k}\right)\left|\sum_{n=0}^{\infty}(2n+1)\sin(\eta_n^S)e^{i\eta_n^S}P_n(\cos\theta)\right|^2_{\theta\cong 180^\circ}, \quad (1)$$

$$\left.\frac{d\sigma_{\text{Cylinder}}}{d\Omega}\right|_{\theta\cong 180^\circ} = \left(\frac{2}{\pi k}\right)\left|\sum_{n=0}^{\infty}\varepsilon_n\sin(\eta_n^C)e^{i\eta_n^C}(\cos n\theta)\right|^2_{\theta\cong 180^\circ}, \quad (2)$$

where k is the wave vector, θ is the angle between **k** and the position vector **r**, $P_n(\cos\theta)$ is the nth order Legendre polynomial, ε_n is the Neumann factor ($\varepsilon_0=1$; $\varepsilon_n=2$, $n>0$), and η_n^S and η_n^C are the phase-shift angles of the nth scattered wave from the sphere and from the cylinder, respectively. Analytical expressions of these phase shift angles are not included here. Their values depend upon the physical properties of the scatterer (density, Poisson's ratio, radius, wave vector of the longitudinal, and the transverse waves of the scatterer material) and the physical properties of the fluid surrounding the scatterer (density, wave vector, and density). Equation (2) holds only for long cylinders (i.e., infinite length), oriented perpendicular to the ultrasound beam axis.

When the size of the scatterers is much smaller than the wavelength, the differential backscatterer cross sections vary as f^3 and f^4 for the cylinder and the sphere, respectively, in agreement with Rayleigh's scattering. Following Wear (1999), at first approximation, we considered that the microarchitecture of trabecular bone could be modeled as long and thin cylindrical scatterers of known diameters, oriented perpendicular to the ultrasound beam axis. In addition, we also computed the backscatter response of an ensemble of equivalent spherical scatterers with an identical diameter. Our bone specimens have been investigated using synchrotron radiation microcomputed tomography (μCT). The diameter of the scatterers was derived from morphometric analysis of the reconstructed 3-D microarchitecture. This will be explained in more detail in a subsequent section (Sec. III D).

The second approach used the formalism based on the autocorrelation function (Ueda and Ichikawa, 1981; Ueda and Ozawa, 1985). This approach has been successfully employed for soft tissue characterization, where the mechanical properties, shape, and size of scatterers are usually unknown (Shung and Thieme, 1993). It assumes an isotropic random statistically homogeneous scattering model, in which weak scattering structures are randomly distributed. This modeling further assumes that the medium is at the focal distance of a circular transducer and that the scatterer's size is smaller than the radius of the transducer, all requirements comfortably met for our experimental procedure. The effects of at-

tenuation in the medium are neglected. Under these conditions the backscatter coefficient relates to the correlation function for the medium properties and is given by the following expression:

$$\mu \propto k^4 \int \int \int_{-\infty}^{\infty} C(x,y,z) \times e^{-j2kz} dx dy dz, \tag{3}$$

where k is the wave vector, and $C(x,y,z)$ is the correlation function which describes the three-dimensional distribution of the specific acoustic impedance inhomogeneities within the medium (according to the x, y, and z axes in space) of the random medium. The z-axis is assumed to be the ultrasound propagation axis. To solve the case of complex biological tissues, in which the exact autocorrelation function is usually unknown, several authors proposed to use equivalent models in which the medium is modeled by simple autocorrelation functions such as Gaussian, exponential, or cylindrical functions structure (Shung and Thieme, 1993). Then, an analytical expression of the backscatterer coefficient can be derived (Romijn *et al.*, 1989). If the characteristic dimension of the scattering structure (correlation length) is very small compared to the wavelength, Rayleigh scattering occurs. In this study, the backscatter coefficient from bone was computed, using the autocorrelation function derived from the 3-D reconstruction of the trabecular bone microarchitecture.

III. MATERIAL AND METHODS

A. Bone specimens

Measurements were performed on 19 human calcanei removed from cadavers of age ranging from 75 to 90 years. Slices of trabecular bone with parallel faces and thickness of approximately 1 cm were cut from the whole calcaneal bone in the sagittal plane, leaving direct access to the highly porous bone (porosity of 90%–95%). The specimens were defatted and refilled with water under vacuum for ultrasonic measurements. The exact sample thickness was measured using an electronic caliper. The experimental protocol of sample preparation has been described in a previous paper from our group (Laugier *et al.*, 1997a).

B. Ultrasonic measurements

Ultrasonic measurements were performed in immersion with a pair of focused broadband transducers with center frequency of 1 MHz and with frequency bandwidth (0.4–1.2 MHz). Both transducers were mounted coaxially, separated by twice the focal length, in a through-transmission normal incidence configuration. The measurement procedure consisted of moving the transducers using a two-dimensional (2D) scanning mechanism over a 60×60 grid (1 mm pixel size, 60×60 mm field of view), and recording the transmitted and reflected raw radio-frequency (rf) signals (Fig. 2). Radio-frequency signals were digitized at 12 MHz. The technical aspects of ultrasonic measurements in transmission have been extensively developed elsewhere (Laugier *et al.*, 1997a) and will not be presented in detail here. Ultrasonic properties in transmission were derived using a conventional substitution technique whereby a pulse transmitted through bone is

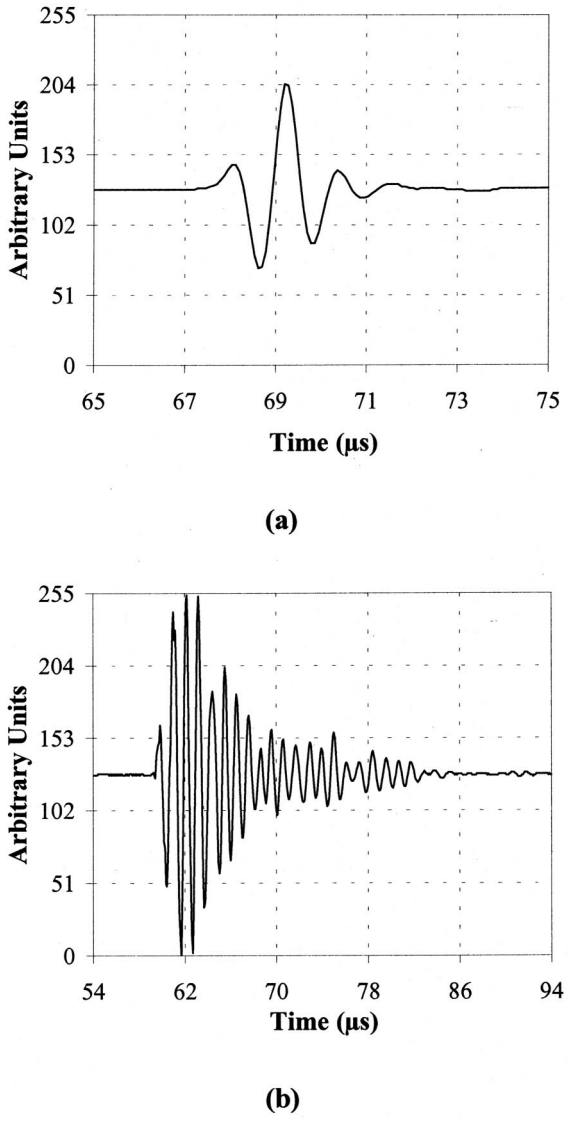


FIG. 2. Signal transmitted through the bone specimen (a) and signal backscattered from the internal bone microarchitecture (b).

Fourier transformed and compared to the Fourier transform of a reference pulse which has been transmitted through water. The attenuation coefficient was derived from the log-ratio of the amplitude spectra. Bone specimens were tested with the ultrasound beam axis parallel to the medio-lateral axis, which corresponds to the clincial measurement configuration. In a previous study, we found that the alignment of the beam axis along the medio-lateral axis of the calcaneus resulted in an incidence angle of 90.7±10.8° with respect to the direction of main orientation of the trabecular network. This indicates that, on the average, the assumption of cylindrical scatterers oriented perpendicularly to the beam axis required in the Faran model was respected (Chaffaï *et al.*, 1999).

Analysis of the reflected signals was performed to determine the backscatter coefficent. A 10.6-μs-gated region set at the focal length of the transducer was used to select the data from a standardized physical location in the bone specimen. This gate location and duration corresponded to a tissue volume of 8 mm in length placed around the center of the

bone specimen. First, the gated signal was multiplied by a windowing function (Hamming) and Fourier transformed to calculate the power spectrum. The backscatter power reflects the frequency dependence of both the instrumentation and the tissue. To deconvolve the response of the measuring equipment from the backscatter data, a calibration spectrum (reference signal) is obtained by reflecting an ultrasonic pulse off a steel plate positioned in water at a distance corresponding to the center of the gated region. The deconvolution of the electromechanical response of the measuring equipment from the backscatter data was computed by performing the log spectral subtraction of the tissue spectra from the calibration spectrum from the standard reflection target as follows:

$$\hat{\mu}_B(f) = 8.68 \ln \frac{\langle S_B(f) \rangle}{S_0(f)}, \tag{4}$$

where $\langle S_B(f) \rangle$ is the spatial average backscattered power spectrum, and $S_0(f)$ the reference spectrum from the standard reflecting target. This yields an apparent backscatter coefficient $\hat{\mu}_B(f)$ in dB. The analysis of the backscattered signal alone requires that there is no contribution of the specular reflection from the sample surface to the signal in the gated signal. In addition, backscatter measurements must be compensated for several sources of error including the signal loss caused by partial transmission at bone interface and attenuation in bone, the frequency dependence of the volume insonified by the transducer, and the effect of the gating function. In a previous work, we have found that 99% of the incident energy was transmitted, and that only 1% of the energy was lost at the interface between the fluid and the fluid-saturated highly porous bone specimen (Chaffaï *et al.*, 2000). This result strongly suggests that the sample surface does not contribute substantially to the scattered signal and that transmission losses at the interface may be neglected. The attenuation was derived from the data acquired in the through-transmit mode and, therefore, could be compensated. The frequency-dependent backscatter coefficient $\mu_B(f)$ was then computed, applying correcting factors compensating for the attenuation and taking into account the geometry of the scattering volume following the method of Roberjot (1996b),

$$\mu_B(f) = \hat{\mu}_B(f) C(f) \text{Factors}(f), \tag{5}$$

where $C(f)$ is the attenuation correction term and Factors(f) is a frequency-dependent scattering volume correction term (i.e., a term accounting for time gate length and frequency-dependent beam width and shape). The attenuation correction term is given by (O'Donnell and Miller, 1981)

$$C(f) = e^{4\hat{\alpha}(f)(z)} \frac{4\hat{\alpha}(f)d}{e^{2\hat{\alpha}(f)d} - e^{-2\hat{\alpha}(f)d}}, \tag{6}$$

where $\hat{\alpha}(f)$ is the estimate of the attenuation coefficient. Following Ueda and Ozawa's work (1985), the volume compensation term may be written

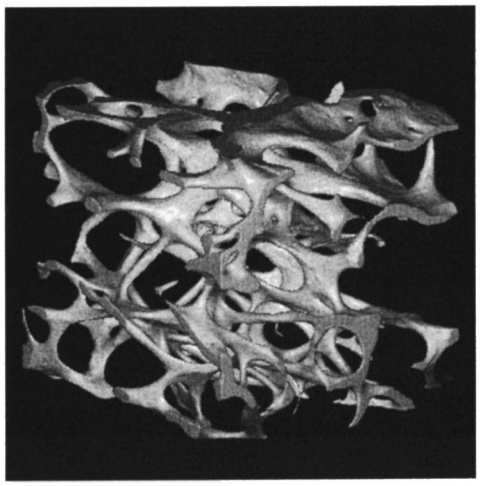


FIG. 3. 3D microarchitecture of a human calcaneus specimen (reconstructed volume of $6.6 \times 6.6 \times 6.6$ mm^3).

$$\text{Factors}(f) = \frac{1}{(0.63)^2} \frac{k^2 a^2}{8\pi d \left[1 + \left(\frac{ka^2}{4F} \right)^2 \right]}, \tag{7}$$

$(1/0.63)^2$: compensation term for the Hamming gate function; d: gate length; z: attenuation path within the bone until the gated volume; $k = 2\pi/\lambda$: wave number; a: transducer radius; F: focal length.

The frequency bandwidth of the 10.6-μs time gate used was much smaller than the frequency range of the useful bandwidth of the transducer. The effect of the convolution of the signal spectra with the Fourier transform of the gate on the frequency dependence of the backscatter coefficient was found to be 1.5% approximately and thus was neglected. When properly compensated for attenuation and geometrical factors, the backscatter coefficient is independent of the measuring instrument, and depends only on the microstructure and acoustic properties of the scattering volume under test (Roberjot *et al.*, 1996b).

The attenuation and the backscatter coefficients were averaged over the entire volume of bone, which were subsequently extracted and investigated using synchrotron radiation microtomography. The site-matching and averaging procedure will be described at the end of the next section.

C. Microstructure assessment

The 3-D microarchitecture of the specimens was investigated using synchrotron radiation μCT at ESRF (European Synchrotron Radiation Facility, Grenoble, France). Indeed, due to the physical properties of synchrotron radiation, this technique provides 3-D images of bone samples at a high spatial resolution (10 μm), with a high signal-to-noise ratio. The acquisition setup is based on 3-D parallel tomography and is described in detail in Salomé *et al.*, 1999. Two-dimensional digitial radiographs of the sample under different angles of view are recorded and further processed by 3-D tomographic reconstruction to get a 3-D image of the sample. The viewing field of radiographs depends on the number and the size of pixels in the digital radiographs. In

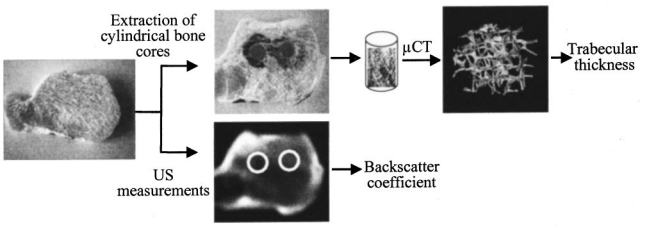


FIG. 4. Illustration of the experimental procedure.

our application, the radiographs are recorded on a 1024 ×1024 charge-coupled-device (CCD) based 2-D detector, and the pixel size is set to 10 μm. This implied that the available field of view is 1 cm^2, and that relatively small specimens could be investigated. Therefore, after initial ultrasonic testing, cylindrical bone cores of diameter 7 mm were cut from the slices, their axis being aligned along the medio-lateral axis (the axis of ultrasonic testing), giving a total of 19 specimens. A volume of interest of 6.6×6.6×6.6 μm^3, centrally located within each cylinder, was reconstructed and analyzed (Fig. 3). The voxel of the reconstructed image is cubic with a 10-μm size in the three directions of space. The mean trabecular thickness for each specimen was derived from the morphometric analysis of the segmented 3-D volume (Parfitt *et al.*, 1983 and Hipp and Simmons, 1997).

Photographs of bone slices were taken after removal of the cylinders. These photographs were used to select circular site-matched regions of interest (ROIs) ultrasonic scans and to measure ultrasonic backscatter at a location corresponding exactly to the extracted cylindrical cores. ROIs were circular, 7 mm in diameter, and contained 44 consecutive pixels. The average value of each parameter was then calculated for each ROI. The whole experimental procedure is summarized in Fig. 4.

D. Modeling

In a first step, each trabecular bone specimen was modeled as a collection of randomly distributed identical cylindrical scatterers (respectively spherical scatterers). For a given bone specimen, scatterers were assumed to have a diameter equal to the mean trabecular thickness derived for the specimen under study. The corresponding frequency dependence of the backscatter coefficient was computed using Eq. (1) [respectively, Eq. (2)]. As the backscatter coefficient was computed for a single scatterer, the level of magnitude was not taken into account in this study. The values of bone characteristics introduced in the model, including the longitudinal velocity (3800 m/s), the shear velocity (2000 m/s), the density (1.95 g/cm^3), and the Poisson's ratio of trabecular frame (0.32) were obtained from Hosokawa and Otani, 1997. Since Hosokawa and Otani's paper analyzes bovine femur, which is likely to have different acoustic properties from human calcaneus, we also performed a parameter study in which the values of the parameters introduced into the Faran model were varied (longitudinal velocity 3500–4000 m/s, shear velocity 1500–2000 m/s, density 1.75–2.0 g/cm^3, and the Poisson's ratio of trabecular frame 0.25–0.35) to determine whether small variations in the acoustic properties have a significant impact on the frequency dependence of backscatter.

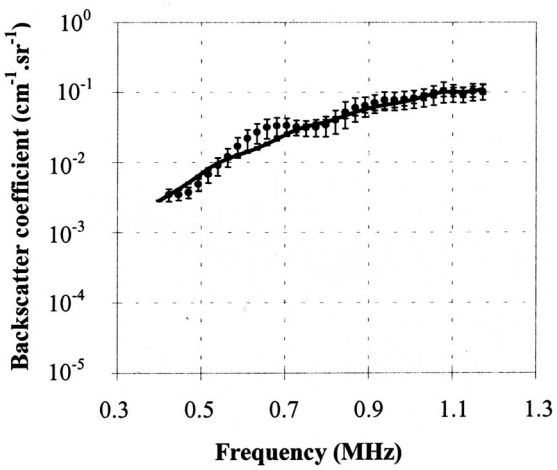


FIG. 5. Average experimental and autocorrelation-based backscatter coefficient for the 19 specimens (● experimental data, — autocorrelation model, error bars = standard error).

In a second step, the backscatter coefficient was modeled using the autocorrelation approach [Eq. (3)]. The autocorrelation function was first numerically computed from data obtained using microtomography data. Ideally, the 3-D autocorrelation function $C(x,y,z)$ was required. However, current computational limitations made it difficult to process the huge amount of data (300 Mbytes/bone specimen). We then decided to simplify the formulation of the autocorrelation function by using the following approximation: $C(x,y,z)=C(x,y)\times C(z)$. This approximation can be partly justified by the anisotropy of the specimens and their orientation in the ultrasonic field. Indeed, the characterization of bone microarchitecture carried out from microtomography measurements using synchrotron radiation showed the direction of main orientation of the trabecular network to be in the sagittal plane, i.e., oriented perpendicular to the ultrasonic beam (z-axis) (Chaffaï *et al.*, 1999). Then, Eq. (3) could be rewritten as

$$\mu \propto k^4 \int \int_{-\infty}^{\infty} C(x,y)dxdy \times \int_{-\infty}^{\infty} C(z)\times e^{-j2z}dz. \qquad (8)$$

This equation shows that the first integral is a frequency-independent term. Given that we were interested in the frequency dependence of the backscatter coefficient, its expression was then derived from the following equation:

$$\mu \propto k^4 \int_{-\infty}^{\infty} C(z)\times e^{-j2kz}dz, \qquad (9)$$

where $C(z)$ is the autocorrelation function in the z-axis direction, and again, the level of magnitude was not taken into account. In order to estimate this autocorrelation function, we used microtomographic slices oriented in the (y,z) plane. First, an autocorrelation function $C_{y_i}(z)$ was estimated for each position y_i along the y-axis. The final autocorrelation $C(z)$ was then deduced by averaging the autocorrelation functions $C_{y_i}(z)$.

IV. RESULTS

Figure 5 shows the experimental measurements of average backscatter coefficient for the 19 cylindrical specimens. Also shown is the autocorrelation-based average backscatter coefficient. The magnitudes of the theoretical curves were arbitrarily adjusted to fit the experimental data. Then, the frequency variation of the backscatter coefficients was characterized by the exponent n of nonlinear power fit least-squares calculations with respect to frequency over the total bandwidth 0.4 to 1.2 MHz. The coefficient of determination (R^2), a measure of the goodness of the fits to the frequency dependence, was used. A good agreement in the frequency variation between experimental and autocorrelation-based backscatter coefficients may be seen. The frequency dependence, was on the average $f^{3.38\pm0.31}$ (R^2 varying from 0.67 to 0.96) for the experimental data and $f^{3.48\pm0.26}$ ($R^2=0.97-0.99$) for the autocorrelation-based backscatter coefficient.

The average values of frequency dependence of the theoretical backscatter coefficients (Faran) were $f^{2.89\pm0.06}$ ($R^2=1$) and $f^{3.91\pm0.04}$ ($R^2=1$) for cylindrical and spherical scatterers, respectively. Here, the standard deviation values for the theoretical values for frequency dependence come from the fact that a different cylindrical (respectively, spherical) model was computed for each bone specimen (i.e., for a given specimen, the scatterer diameter is the estimated mean trabecular thickness). In the parameter study, the variation in longitudinal velocity, shear velocity, density, and Poisson's ratio varied the exponent n by less than 2%, which is much less than the difference in experimental and theoretical frequency dependences, suggesting that the disparity between our measurements and the cylindrical model (or the spherical model) cannot be accounted for by inappropriateness of the selected acoustical values.

Typical examples of experimental and autocorrelation-based backscatter coefficients are illustrated in Fig. 6 for samples 1 and 3. For sample 1, the frequency dependence is $f^{2.76}$ and $f^{2.92}$ for experimental data and the autocorrelation-based model, respectively, for sample 3: the frequency dependence is $f^{3.86}$ and $f^{3.83}$ for experimental data and the autocorrelation-based model, respectively. In addition, it may be seen that the frequency dependence is close to that of the backscatter coefficient from a small cylinder for specimen 1 and that of the backscatter coefficient from a small sphere for specimen 3.

Descriptive statistics of the mean trabecular thickness and of the exponent of the experimental and autocorrelation-based backscatter coefficient are summarized in Table I. No statistically significant difference between the experimental exponent and autocorrelation-based model exponent was found (Wilcoxon paired test, $p>0.05$). On the other hand, the correlation coefficient between the experimental exponent and autocorrelation-based model exponent values was highly significant but moderate ($r=0.55$, $p<0.05$).

V. DISCUSSION

A new model of backscatter from human trabecular bone based on the autocorrelation approach was presented in this study and compared to experimental data and analytical solutions of scattering from small cylindrical and spherical scatterers. The frequency dependences for experimental and autocorrelation-based backscatter coefficients were $f^{3.38\pm0.31}$ and $f^{3.48\pm0.26}$. It is important to note that there is a good agreement between both experimental data and the autocorrelation-based model, and this, in spite of the rather restrictive assumption of weak scattering. In addition, we keep in mind that although autocorrelation models apply well in soft tissues, they describe the scattering from solid scatterers less completely. In particular, this model does not account for shear waves in the scatterers. In our study, no significant difference has been found between the experimental and the modeled values of the exponent n.

The study, following Wear's work (1999), also investigates the cylindrical Faran model and, in addition, the spherical Faran model. The results based on Faran theory show qualitative agreement with experimental data and quite close agreement with the work of Wear (1999). Wear has found for 16 human calcaneal specimens a good agreement in the frequency variation between experimental and theoretical

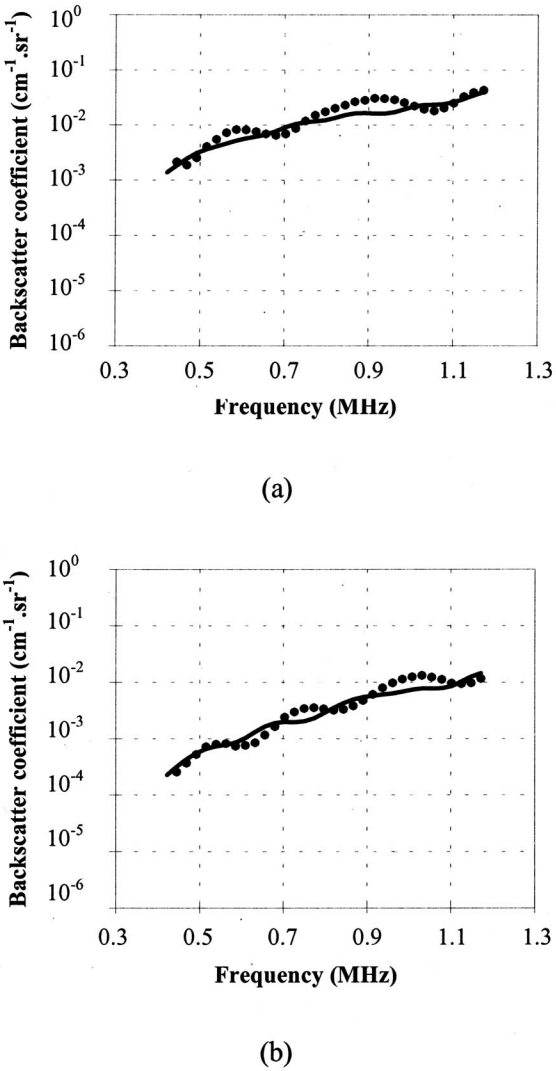


FIG. 6. Examples of the experimental and the autorrelation-based model backscatter coefficient: (a) specimen 1, (b) specimen 3. (● experimental data, — autocorrelation model.)

TABLE I. Descriptive statistics of the mean trabecular thickness and the exponent of the experimental, Faran models, and autocorrelation-based backscatter coefficients.

Specimen	Mean trabecular thickness (μm)	Exponent n			
		Faran cylindrical model	Faran spherical model	Experimental	Autocorrelation
1	53	2.937	3.956	2.76	2.9
2	54	2.935	3.955	3.23	3.49
3	43	2.957	3.971	3.86	3.83
4	44	2.955	3.970	3.68	3.53
5	53	2.937	3.956	3.21	3.34
6	82	2.859	3.896	3.83	3.63
7	92	2.823	3.868	3.77	3.74
8	59	2.924	3.946	3.15	3.84
9	67	2.900	3.931	3.12	3.53
10	76	2.878	3.910	3.34	3.59
11	107	2.766	3.822	3.75	3.57
12	95	2.812	3.858	3.54	3.22
13	71	2.893	3.922	3.05	3.55
14	90	2.832	3.875	3.03	3.39
15	72	2.889	3.916	3.25	3.26
16	82	2.857	3.894	3.24	2.98
17	82	2.857	3.894	3.48	3.31
18	72	2.890	3.921	3.67	3.62
19	53	2.930	3.951	3.27	3.75

(Faran theory-based approach) backscatter coefficients and found experimentally a frequency dependence in $f^{3.2\pm0.2}$ close to the variation in f^3 obtained for cylindrical scatterers (Wear, 1999). Whether our experimental data showing a frequency dependence in $f^{3.38\pm0.31}$ are significantly different or not from those reported by Wear remains unsolved. Differences in methods make comparison somewhat difficult: these include different bone specimens, frequency bandwidth used in calculations, and differences in data analysis.

Given the low frequencies falling between 0.4 and 1.2 MHz (wavelength ranging from 1.25 to 3.75 mm) used in the study and given the mean trabeculae thickness falling in the range 43–107 μm, the trabeculae should behave approximately as Rayleigh scatterers. However, a more exact evaluation of ka (k wave vector and a trabecular thickness) yields values between 0.07 and 0.5, suggesting that Rayleigh scattering ($ka \ll 1$) must be considered only as a crude approximation. The numerical computation of the Faran solution for cylinders and spheres (with diameters equal to the mean trabecular thickness) yielded frequency dependences in $f^{2.89\pm0.06}$ and $f^{3.91\pm0.04}$, respectively, which appears to be quite close to Rayleigh scattering for cylinders (f^3) and spheres (f^4). As expected, we found that the smaller the trabeculae the closer to Rayleigh scattering the frequency dependence of the Faran cylindrical and spherical models (Table I).

In contrast, experimental and autocorrelation data showed substantial between-sample variation (experimental frequency dependence ranging from 2.76 to 3.86, autocorrelation frequency dependence ranging from 2.90 to 3.84). The frequency dependence was closer to the cylindrical model for some specimens and closer to the spherical model for some other specimens, but these values were somewhat counterintuitive since no evidence of a relationship between trabecular thickness and frequency dependence of scattering could be found. In a simple model with identical scatterers, e.g., spherical or cylindrical scatterers, one would expect that the larger trabeculae would have a lower power law. It may be that, due to the tremendous variation in microarchitecture encountered in human trabecular bone, the frequency dependence of the backscatter coefficient from trabecular bone cannot be described accurately by a single scattering model, i.e., neither by scattering from a collection of identical cylindrical scatterers, nor by spherical scatterers. Different models (or combination of models) might be needed to fully explain the backscatter from bone specimens with different microarchitecture, especially given the fact that we know that the shape of the trabeculae tends to evolve from a platelike to a rodlike shape during aging. Future studies are warranted to determine whether more realistic approaches are needed, like scattering by short cylinders or ellipsoidal structures, to obtain a more exact prediction of scattering from the trabecular network.

The interesting observation of this study, according to which scattering by the trabecular structure can be modeled using the autocorrelation functions, opens interesting new routes toward the understanding of the interaction between ultrasonic waves and trabecular bone. Using image processing tools, thinning or thickening the trabeculae, disconnecting them, or even completely removing some of them becomes possible. A large variety of different bone structures could be obtained while modifying the reconstructed 3-D microarchitectures. That would provide a useful insight into the relationship of bone microarchitecture and ultrasonic scattering.

The observation of a frequency dependence in $f^{3.38\pm0.31}$ for ultrasound backscatter reinforces the findings of Wear and their implication on the respective role of absorption and scattering phenomena in the attenuation mechanisms of the ultrasound wave traveling through trabecular bone. Indeed,

in a previous study by our group, the frequency dependence of the attenuation coefficient of the same collection of bone specimens has been documented and reported to vary like $f^{1.2\pm0.2}$ (Chaffaï *et al.*, 2000). The combination of these two results brings us to the conclusion that scattering probably gives a minor contribution to the ultrasonic attenuation by bone structure.

Although the experimental measurements are closer to the autocorrelation model than the cylindrical or the spherical model, the numerous error sources, precision limitations, and numerous simplifying assumptions for both autocorrelation and Faran models makes difficult the quantitative assessment of the level of agreement of each model with experimental data. The difference between the Faran model and measurements is not that dramatic given the accuracy and limitation of the experiment. The Faran model seems to be reasonably accurate. Furthermore, in spite of a good agreement in the averaged frequency dependence of experimental and autocorrelation-based data, the correlation between experimental and modeled values remained moderate ($r=0.55$, $p<0.05$). These results underline the limits of this study. Several sources of potential errors may explain the absence accuracy and of a stronger correlation. On one hand, backscatter measurements are difficult, and the level of accuracy in experimental measurements is limited as illustrated by a recent interlaboratory comparison of measurements of ultrasonic properties on tissue-mimicking samples (Madsen *et al.*, 1999). We have previously demonstrated good agreement between theory and experiment for this measurement for suspended spheres in gelatin over a very broad bandwidth (2 to 60 MHz) (Roberjot *et al.*, 1996b). The current measurements in bone, however, are over a small bandwidth, and the level of accuracy remains poorly qualified under such experimental conditions. On the other hand, the fact that the autocorrelation function $C(x,y,z)$ was approximated by $C(x,y)\times C(z)$ and obtained using a single tomographic slice instead of using the whole 3-D data set could potentially bias the estimates. The accuracy was limited by the degree of inhomogeneity of the volume under investigation. A more accurate estimate of the autocorrelation function would require computing the whole 3-D data set from $(660)^3$ voxels (equivalent to 300 Mbytes), which remains a difficult task with currently available computing facilities at our center. The calculation of the 3-D autocorrelation function would allow a more detailed analysis of the impact of bone microarchitecture, particularly that of structural anisotropy, on backscatter. A third potential technical source of error is associated with statistical noise. Backscatter measurements were spatially averaged over a region of 7 mm in diameter, containing 44 consecutive lines in steps of 1 mm in both x and y directions. Averaging was necessary to remove pure statistical variations in the reflected signal due to the random phase of the elementary waves scattered by each scattering trabeculae. It may be that, for some bone specimens, interference noise was not totally removed (e.g., by averaging a limited number of partially correlated lines), resulting in significant residual spectral noise. This effect may decrease the accuracy of experimental measurements of the exponent n of the power fit. Measurements achieved while using higher frequencies or considering larger regions of interest would contribute to decrease spurious signal interference and would assure a more accurate estimation. Finally, several simplifying assumptions represent the candidate that could potentially affect the exponent of frequency dependence by a few tenths or more. Both models assume an isotropic medium and ignore the impact of the coherent component of scattering, phase cancellation, or multiple scattering. The results obtained in this study are based on the assumptions that phase-cancellation effects are insignificant (Strelitzki *et al.*, 1999). One of the problems is that the trabeculae cannot be regarded as weak scatterers, and in practice, the assumption of weak scattering may not be strictly satisfied and multiple scattering may play a role. When a wave undergoes multiple scattering as it traverses a strongly scattering medium, the transmitted wave can be divided into a coherent term which corresponds to the attenuated incident wave and an incoherent wave which contains the contribution to the transmitted field of all the multiply scattered waves. As a consequence, the temporal profile of the pulse transmitted through the medium is spread over a long time interval compared to the incident pulse time profile. Considering experimental results derived from ultrasound transmission through bone at frequencies less than 1 MHz (the clinically useful frequencies), no broadening of the transmitted pulse temporal profile, apart from the effect of the frequency-dependent attenuation, has been evidenced (Fig. 2). Furthermore, it has been argued that, would multiple scattering be present, its impact on measurements could be decreased dramatically by the strong attenuation encountered in trabecular bone (Wear, 1999). For the time being, experiments do not suggest the presence of strong multiple scattering. Future experimental studies should nevertheless be carried out to specifically address the issue of multiple scattering in cancellous bone.

VI. CONCLUSION

We have tested several theoretical models of ultrasound backscatter from trabecular bone. The experimental frequency dependence in $f^{3.38\pm0.31}$ is close to the autocorrelation model in $f^{3.48\pm0.26}$ and qualitative agreement was found with analytical Faran scattering theory by cylindrical ($\sim f^3$) objects. Future studies are warranted to determine whether more realistic approaches are needed, like scattering by short cylinders or ellipsoidal structures, to obtain a more exact predicition of scattering from the trabecular network. The good prediction obtained by modeling the backscatter coefficient using the autocorrelation function of the medium opens interesting prospects for the investigation of the influence of bone microarchitecture on ultrasonic scattering.

Bauer, D. C., Glüer, C. C., Cauley, J. A., Vogt, T. M., Ensrud, K. E., Genant, H. K., and Black, D. M. (**1997**). ''Broadband ultrasound attenuation predicts fractures strongly and independently of densitometry in older women,'' Arch. Int. Med. **157**, 629–634.

Benhamou, C. L., Lespessailles, E., Jacquet, G., Harba, R., Jennane, R., Loussot, T., Tourliere, D., and Ohley, W. (**1994**). ''Fractal organization of trabecular bone images on calcaneus radiograph,'' J. Bone Miner. Res. **9**, 1909–1917.

Chaffaï, S., Padilla, F., Berger, G., and Laugier, P. (**2000**). ''*In vitro* measurement of the frequency dependent attenuation in cancellous bone between 0.2–2.0 MHz,'' J. Acoust. Soc. Am. **108**, 1281–1289.

Chaffaï, S., Peyrin, F., Berger, G., and Laugier, P. (**1999**). ''Variation of human cancellous bone ultrasonic properties with density and microarchitecture,'' in IEEE Ultrasonics Symposium Proceedings, edited by S. C. Schneider, M. Levy, and B. R. McAvoy (Institute of Electrical and Electronic Engineering, New York), pp. 1303–1306.

Faran, J. J. (**1951**). ''Sound scattering by solid cylinders and spheres,'' J. Acoust. Soc. Am. **23**, 405–418.

Glüer, C. C. (**1997**). ''Quantitative ultrasound technique for the assessment of osteoporosis: Expert agreement on current status,'' J. Bone Miner. Res. **12**, 1280–1288.

Gordon, C. L., Webber, C. E., Christoforou, N., and Nahmias, C. (**1997**). ''*In vivo* assessment of trabecular bone structure at the distal radius from high-resolution magnetic resonance images,'' Med. Phys. **24**, 585–593.

Hans, D., Dargent-Moline, P., Schott, A. M., Sebert, J. L., Cormier, C., Kotski, P. O., Delmas, P. D., Pouilles, J. M., Breart, G., and Meunier, P. J. (**1996**). ''Ultrasonographic heel measurements to predict hip fracture in elderly women: The Epidos prospective study,'' Langmuir **348**, 511–514.

Hans, D., Wu, C., Njeh, C. F., Zhao, S., Augat, P., Newitt, D., Link, T., Lu, Y., Majumdar, S., and Genant, H. K. (**1999**). ''Ultrasound velocity of trabecular cubes reflects mainly bone density and elasticity,'' Calcif. Tissue Int. **64**, 18–23.

Hipp, J. A., and Simmons, C. A. (**1997**). ''Method-based differences in the automated analysis of the three-dimensional morphology of trabecular bone,'' J. Bone Miner. Res. **12**, 942–947.

Hosokawa, A., and Otani, T. (**1997**). ''Ultrasonic wave propagation in bovine cancellous bone,'' J. Acoust. Soc. Am. **101**, 558–562.

Insana, M. F., Wagner, R. F., Brown, D. G., and Hall, T. J. (**1990**). ''Describing small-scale structure in random media using pulse-echo ultrasound,'' J. Acoust. Soc. Am. **87**, 179–182.

Laugier, P., Droin, P., Laval-Jeantet, A. M., and Berger, G. (**1997a**). ''*In vitro* assessment of the relationship between acoustic properties and bone mass density of the calcaneus by comparison of ultrasound parametric imaging and quantitative computed tomography,'' Bone **20**, 157–165.

Laugier, P., Giat, P., Chappard, C., Roux, C., and Berger, G. (**1997b**). ''Clinical assessment of the backscatter coefficient in osteoporosis,'' in IEEE Ultrasonics Symposium Proceedings, edited by S. C. Schneider, M. Levy, and B. R. McAvoy (Institute of Electrical and Electronic Engineers, New York), pp. 1101–1103.

Madsen, E. L., Dong, F., Frank, G. R., Garra, B. S., Wear, K. A., Wilson, T., Zagzebski, J. A., Miller, H. L., Shung, K., Wang, S. H., Feleppa, E. J., Liu, T., O'Brien, W. D., Toop, K. A., Sanghvi, N. T., Zaitsev, A. V., Hall, T. J., Fowlkes, J. B., Kripfgans, O. D., and Miller, J. G. (**1999**). ''Interlaboratory comparison of ultrasonic backscatter, attenuation, and speed measurements,'' J. Ultrasound Med. **18**, 615–631.

Majumdar, S., Newitt, D., Mahtur, A., Osman, D., Gies, A., Chiu, E., Lotz, J., Kinney, J., and Genant, H. K. (**1996**). ''Magnetic resonance imaging of trabecular bone structure in the distal radius: Relationship with X-ray tomographic microscopy and biomechanics,'' Osteoporosis Int. **6**, 376–385.

Müller, R., Häuselmann, H. J., Hildebrand, T., and Rügsegger, P. (**1996**). ''*In vivo* reproducibility of three-dimensional structural properties of noninvasive bone biopsies using 3D-pQCT,'' J. Bone Miner. Res. **11**, 1745–1750.

Nicholson, P. H., Müller, R., Lowet, G., Cheng, X. G., Hildebrand, T., Rüegseger, P., van der Perre, G., Dequeker, J., and Boonen, S. (**1998**). ''Do quantitative ultrasound measurements reflect structure independently of density in human vertebral cancellous bone?,'' Bone **23**, 425–431.

O'Donnell, M., and Miller, J. G. (**1981**). ''Quantitative broadband ultrasonic backscatter coefficient: An approach to nondestructive evaluation in acoustically inhomogeneous materials,'' J. Appl. Phys. **52**, 1056–1065.

Parfitt, A. M., Mathews, C. H. E., Villanueva, A. R., and Kleerekoper, M. (**1983**). ''Relationships between surface, volume and thickness of iliac trabecular bone in aging and in osteoporosis: Implications for the microanatomic and cellular mechanisms of bone loss,'' J. Clin. Invest. **72**, 1396–1409.

Roberjot, V., Bridal, L., Laugier, P., and Berger, G. (**1996b**). ''Absolute backscatter coefficient over a wide range of frequencies in a tissue-mimicking phantom containing two populations of scatterers,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 970–978.

Roberjot, V., Laugier, P., Droin, P., and Berger, G. (**1996a**). ''Measurement of integrated backscatter coefficient of trabecular bone,'' IEEE Ultrasonics Symposium, edited by S. C. Schneider, M. Levy, and B. R. McAvoy (Institute of Electrical and Electronic Engineers, New York), pp. 1123–1126.

Romijn, R. L., Thijssen, J. M., Van Delft, J. L., De Wolff-Rouendaal, D., Van Best, J., and Oosterhuis, J. A. (**1989**). ''*In vivo* ultrasound backscattering estimation for tumor diagnosis: An animal study,'' Ultrasound Med. Biol. **15**, 471–479.

Roux, C., Roberjot, V., Kolta, S., Porcher, R., Fechtenbaum, J., and Laugier, P. (**1999**). ''Broadband ultrasonic backscattering (BUB) measurements at the oscalsis in osteoporosis,'' ASBMR 21st Annual Meeting, St. Louis, Missouri [J. Bone Miner. Res. **14**, S375 (1999)].

Salomé, F., Peyrin, F., Cloetens, P., Odet, C., Laval-Jeantet, A. M., Baruchel, J., and Spanne, P. (**1999**). ''A synchrotron radiation micromography system for the analysis of trabecular bone samples,'' Med. Phys. **26**, 2194–2204.

Strelitzki, R., Metcalfe, S. C., Nicholson, P. H. F., Evans, J. A., and Paech, V. (**1999**). ''On the ultrasonic attenuation and its frequency dependence in the oscalcis assessed with a multielement receiver,'' Ultrasound Med. Biol. **25**, 133–141.

Shung, K. K., and Thieme, G. A. (**1993**). *Ultrasonic Scattering in Biological Tissues* (CRC, Boca Raton).

Thomson, P. W., Taylor, J., Oliver, R., and Fisher, A. (**1998**). ''Quantitative ultrasound (QUS) of the heel predicts wrist and osteoporosis-related fractures in women age 45–75 years,'' J. Clin. Densitometry **1**, 219–225.

Ueda, M., and Ichikawa, H. (**1981**). ''Analysis of an echo signal reflected from a weakly scattering volume by a discrete model of the medium,'' J. Acoust. Soc. Am. **70**, 1768–1775.

Ueda, M., and Ozawa, Y. (**1985**). ''Spectral analysis of echoes for backscattering coefficient measurement,'' J. Acoust. Soc. Am. **77**, 38–47.

Wear, K. A., and Garra, B. S. (**1998**). ''Assessment of bone density using ultrasonic backscattering,'' Ultrasound Med. Biol. **24**, 689–695.

Wear, K. (**1999**). ''Frequency dependence of ultrasonic backscatter from human trabecular bone: Theory and experiment,'' J. Acoust. Soc. Am. **106**, 3659–3664.

The role of leaf structure in vibration propagation

C. Magal[a)]
Université de Tours, Institut de Recherche sur la Biologie de l'Insecte, IRBI CNRS ESA 6035, 37000 Tours, France

M. Schöller and J. Tautz[b)]
Universitaet Wuerzburg, Biozentrum, 931/888-4309, Lehrstuhl fuer Verhaltenphysiologie und Soziobiologie, am Hubland, D-97074 Wuerzburg, Germany

J. Casas[c)]
Université de Tours, Institut de Recherche sur la Biologie de l'Insecte, IRBI CNRS ESA 6035, 37000 Tours, France

(Received 23 August 1999; accepted for publication 31 March 2000)

The leaf and its structural components play a key role in the propagation of short transient signals produced by insects. In this paper, it is shown how the complex structure of an apple leaf could be modeled by a much simpler one for the analysis of vibratory signal propagation. Waves were produced by impacts of small spheres and the propagation studied using two laser vibrometers, followed by a wavelets analysis. Three components of the leaf were investigated: the midvein, minor veins, and the interspaced homogeneous regions making up the leaf lamina. The loss of signal energy over the leaf lamina and across minor veins and midvein was studied. For the midvein, the loss of energy decreased from 80% at the leaf base to 40% at the apex. For minor veins, the loss of energy decreased from 70% at the leaf base to 31% at the apex. The loss in homogeneous regions was 40%. A signal decomposition into two frequency ranges, above and below 1.7 kHz, showed that the midvein acted as a low-pass filter. As energy loss was mainly a function of vein diameter and not vein type, veins smaller or equal to 0.2 mm were considered as equivalent to homogeneous regions. Hence, a model leaf reduced to the leaf lamina and veins with a diameter >0.2 mm is retained for the study of signal propagation in a leaf. [S0001-4966(00)02107-X]

PACS numbers: 43.80.Jz, 43.80.Ka, 43.80.Lb [WA]

INTRODUCTION

Vibrations in plant leaves play an important role in many different ecological processes ranging from arthropod communication and spore liberation of fungi to gas exchange with the atmosphere (Ewing, 1989; Jones, 1992; Henry, 1994; Schuepp, 1993; Mulkey *et al.*, 1996; Stewart, 1997; Smith *et al.*, 1997). Hence, knowledge about wave propagation and vibrations in plants can not only help understand the above ecological processes, but can also have applied implications. In particular, the prediction of expected changes in boundary layers around leaves and transpiration rate due to increased air turbulence can aid in determining the expected response of vegetation to climatic change (Sellers *et al.*, 1996).

Many arthropods produce and use vibratory signals in leaves. Spiders and several insect species use leaves as a sound-radiating surface in order to increase their acoustical visibility in sexual communication (reviewed in Barth, 1998). Some spiders and predatory insects make use of the vibrations produced by their prey while walking on plants to detect and attack them. A very large group of organisms, the parasitic wasps, is able to localize and parasitize larvae of phytophagous insects living in plant materials (leaves, fruits, logs) by using the vibrations produced by the larvae when moving or eating (reviewed by Meyhöfer and Casas, 1999). We recently showed that lepidopteran larvae living inside a tunnel built within the depth of an apple leaf (called a mine) are also able to perceive and respond to vibrations produced by the hunting wasps and to behave accordingly (Meyhöfer *et al.*, 1997; Bacher *et al.*, 1997). The vibrations produced by the prey and the wasps are invariably short-term transients, almost impact-like signals lasting from 10 to 700 ms, with peaks of maximal intensity located between a few hundred Hz and 16 kHz, depending on the behavior of the insect (Meyhöfer *et al.*, 1994; Bacher *et al.*, 1996). In this case both antagonists are producing and reacting to vibrations, leading to a coupled system. We do not know which characteristics of the vibrations these animals really use. However, work on other insects and spiders has shown that most characteristics of a signal can be used, such as the amplitude, the maximal amplitude, the phase shift, the frequency content, etc. (reviewed in Barth, 1998). In all cases, the leaf acts as a transmission channel of the vibratory information, and a good knowledge of the signal propagation is therefore a requisite for further understanding systems such as those described above.

Few studies have been devoted to the analysis of vibrations in leaves, perhaps because of the difficulties associated with their delicate structures and the inherent need to use

[a)]Electronic mail: magal@univ-tours.fr
[b)]Electronic mail: tautz@biozentrum.uni-wuerzburg.de
[c)]Electronic mail: casas@univ-tours.fr

contact-free methods whenever possible (Martens and Michelsen, 1981; Tang *et al.*, 1986; Martens, 1990). The most complete study so far used laser doppler vibrometry of leaves in the context of noise abatement (Martens, 1990). These authors showed clearly that several theoretical models can be used as good approximations for clamped leaves. However, they also showed that none of the available models was sufficiently good for understanding naturally hanging leaves, which are too complex and require a new approach, unexplored so far.

The aim of this work was to characterize the propagation of transient vibrations in unclamped apple leaves. In particular, we aimed to identify which of the structures of the leaf must be retained in order to predict the transmission of information in the prey–predator interaction described above. Given the lack of well-founded theory of wave propagation in such structures/materials, we opted for a purely experimental approach followed by a signal analysis. Given the highly complex structure of a mine, we opted for the simpler situation of an unmined leaf.

We use impacts made by a small metal ball as a good approximation to real signals (following a preliminary study by Casas *et al.*, 1998) and a combination of two laser Doppler vibrometers. The study of nonstationary signals, where transient events appear that cannot be predicted, necessitates techniques different from Fourier analysis (Meyer, 1993). Hence, we characterize the signals in two steps. First, we use a simple and well-known criterion in signal analysis, the energy content, and measure its loss through the different structures (Cohen, 1995). Then, we refine the analysis by focusing on the frequency-dependent losses. To do so, we decompose the signals using wavelets Daubechies of order 2, because this type of wavelet is particularly well-adapted for the localization of sharp discontinuities (Misiti *et al.*, 1996). Then, we analyze the loss of energy as a function of the frequency range. We discuss the results in the light of prey–predator systems and in more general terms.

I. MATERIALS AND METHODS

The setup was composed of two OMETRON laser vibrometers VS 100, with a sensitivity of 1 μm/s and a maximal frequency range of 50 kHz, placed on an antivibration table, an unmined apple leaf, a small metal ball weighing 2.5 mg, a Siglab data recorder (Digital Signal Process Technology) working at a sampling frequency of 50 kHz per channel, and a computer. The apple leaf was maintained horizontally using the setup of Meyhöfer *et al.* (1994). The petiole was placed through a hole in the lid of a glass tube filled with water. Plasticine was used to prevent any further movement of the petiole. An unwrapped section of the petiole (at least 5 mm long) was left between the plasticine and the leaf lamina. We distinguished three structural parts of a leaf: the midvein, the minor veins radiating out of the midvein, and the homogeneous regions (Fig. 1). A homogeneous region, also called a sector, was a region of the lamina between two minor veins. In reality, these regions are made of several different tissues and are homogeneous only at a macroscopic scale.

An electrical magnet was used to drop the metal ball on specific locations on the leaf from a height of 2 cm, leading to a calculated potential energy at impact of 5×10^{-7} Nm. The impact point was always in a sector, which was the location where the prey lives. The distance between the impact point and the first measurement point was 2 mm, the smallest possible distance. The distance between the two measurement points was 4 mm. Longer distances were not possible without measuring reflections from minor veins or from the major vein. A distance of 5 mm from the border was kept to avoid undue effects from reflections of waves from the border.

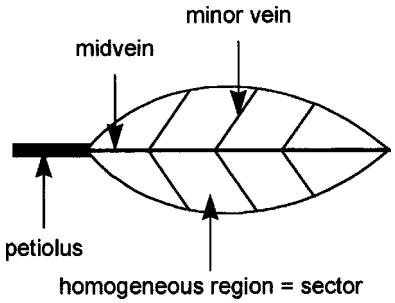


FIG. 1. Description of an apple leaf. Sectors and minor veins are numbered from base (near petiole) to apex.

Several leaves were used, as their state changed rapidly due to desiccation. Leaves were taken from young seedlings and weighed 125 mg on average (dry weight, s.d.=38.6 mg, $n=7$). The length was 5.9 cm on average (s.d.=0.63, $n=7$) and the width was 3.71 cm on average (s.d.=0.28, $n=7$). Signals were recorded for a period of 5.85 ms (300 data points). The first sampled point was always set at the same instant before the impact. The energy was calculated as

$$E=\sum_{i=1}^{300} x(t)^2,$$

with $x(t)$ being the velocity amplitude of the signal at time t. Preliminary experiments showed that at least 90% of the total energy of the signals was contained within the 5.85-ms window. This translated into a very small difference (ca. 1%) when estimating the energy losses between two measurement points, as the tails of the two signals were almost identical. Signals were analyzed using wavelets energy decomposition analysis (Misiti *et al.*, 1996). We decomposed each signal into five frequency ranges, rather than into the eight frequency levels given by the wavelet decomposition, because it gave a fine-enough frequency decomposition. We then calculated the proportion of the total energy contained in each frequency range. In order to avoid analyzing signals which were beyond the sensitivity threshold of the animals, we did a similar analysis on signals from wriggling prey pupae which were known to trigger behavioral responses in the wasps. In all cases reported here, the energy content of the signals at the second measurement point was at least one order of magnitude higher than the energy content of the weakest signals from a pupa. As done in the analysis of vibration signals (Wahl and Bolton, 1993), we presented the results of the decomposition in frequency ranges rather than in a wavelet scales.

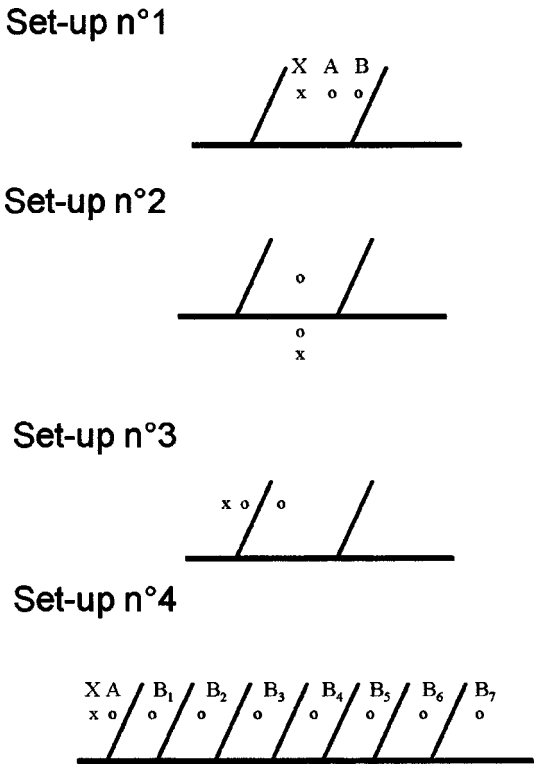


FIG. 2. Description of the measurement setups. *X* is the impact point, *A* is the first measurement point, always located near the impact and used to estimate the input to the system, and *B* is the second measurement point. In the last setup, B_i refers to the different locations of the second measurement point.

II. RESULTS

We first analyzed the role of structural elements and their geometry in determining the loss of total energy between the first and second measurement points. We then considered frequency-dependent energy losses through the same structural elements.

A. The role of structural elements on the loss of total energy

For the ease of presentation, we equated the source location (the impact point) and the first measurement. The second measurement point was then referred to as ''the receiver.''

The loss of energy within a sector was estimated by locating the two measurement points and the impact point within one sector (Fig. 2, setup no. 1). When the source of the signal and the receiver were located on the same sector, the average percentage of energy loss was 44% (s.d.=12%, 42 impacts in 11 homogeneous regions on three leaves). There was no significant difference between sectors near the stem and at the tip of the leaf.

The loss of energy through the midvein was estimated by locating the two measuring points on each side of it (Fig. 2, setup no. 2). Experiments were conducted on two leaves (six sectors in the first leaf and ten in the second, 39 measurements total). When the midvein was located between the source and the receiver, the percentage of energy lost through the midvein varied along the midvein (Fig. 3). The loss was 80% near the stem and decreased linearly to 40% at the apex.

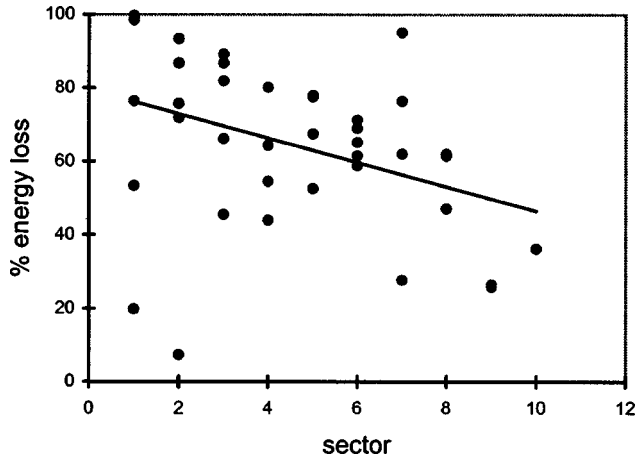


FIG. 3. Percentage of energy loss through the midvein as function of the position of the sector. Sector number 1 is near the petiole and sector number 10 is at the apex of the leaf. Black points are measurements and the straight line was obtained by linear regression ($y=79-3.3x$, $t=-2.47$, $P=0.0195$). Leaving the two outliers near the origin gives $y=90.28-4.9x$, $t=-4.97$, $P=0.0000$.

The loss of energy through the minor veins was estimated by locating the two measurement points on each side of a minor vein (Fig. 2, setup no. 3). The loss of total energy between the source and the receiver was on average 49% ($n=108$ impacts on four leaves, s.d.=19%). The loss of energy through a minor vein was 70% near the petiole and 31% near the apex (Fig. 4). There are no gradient of energy loss along a minor vein.

The dependency of energy loss on the position of the minor veins may have an important effect on the energy propagation through the whole leaf. In order to study the cumulative effect of several minor veins, we predicted the loss of energy through the whole leaf under two different assumptions, the first using the calculated relationship between energy loss and vein position (Fig. 4), and the second

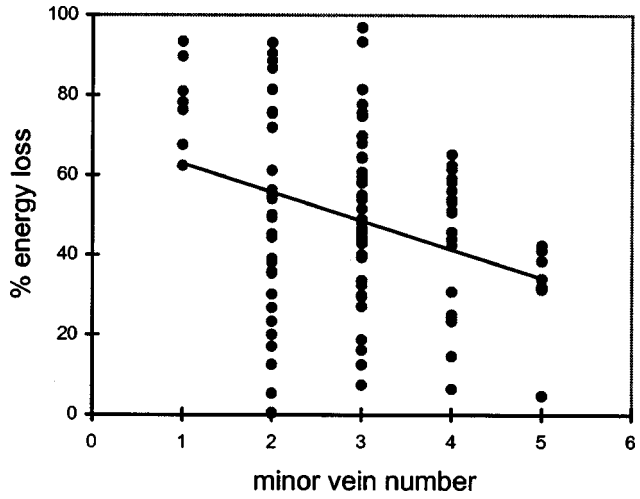


FIG. 4. Percentage of energy loss through a minor vein as function of its position. Minor vein number 1 is near the petiole, irrespective of its location on the side of the midvein, and minor vein number 5 is at the apex of the leaf. Black points are measurements and the straight line was obtained by linear regression ($y=70-7.2x$, $t=-3.51$, $P=0.0007$).

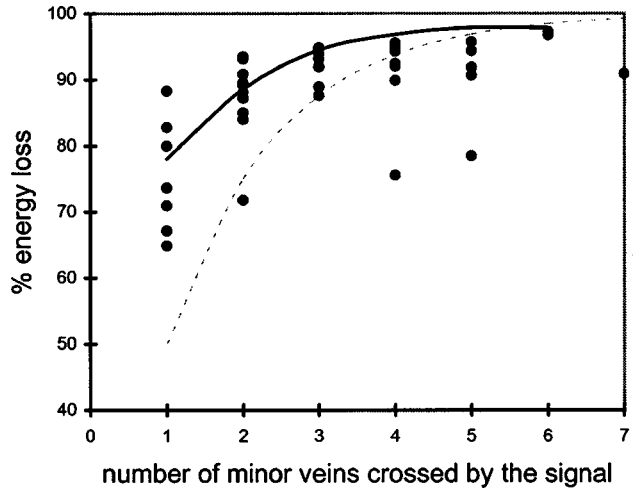


FIG. 5. Percentage of energy loss as function of number of minor veins crossed by the signal. Black points are measurements. The dotted line represents predicted values under the assumption of a constant energy loss of 49%. The black line represents predicted values using observed values of the position-dependent energy loss given in Fig. 4.

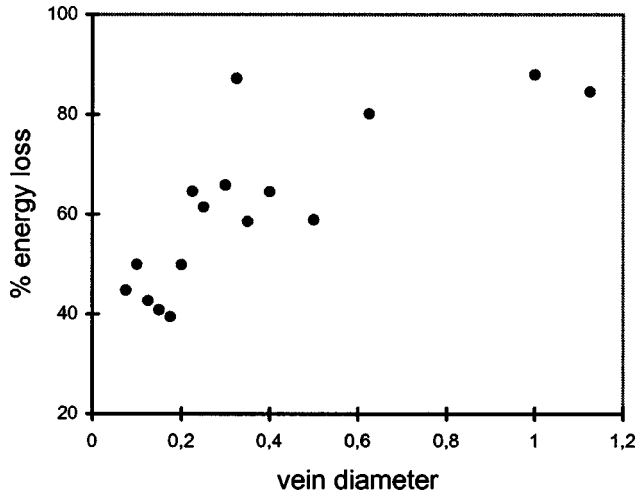


FIG. 6. Average percentage of energy loss through midveins and minor veins as function of vein diameter.

using a position-independent mean value of energy loss (49%). We then measured experimentally the loss on two leaves using a new setup (Fig. 2, setup no. 4, 12 measurements on one leaf and 29 measurements on a second leaf). The impact point and the measurement point *A* were, as before, in the same homogeneous region. The measurement point *B* was located on the other side of a minor vein. In each successive experiment, the number of crossed minor veins was incremented. There was a maximum of seven minor veins between points *A* and *B*. The increasing distance between the two measurement points implied that the signal measured in *B* incorporated the effects of multiple reflections from borders and veins. Hence, our setup cannot identify the sole distance effect.

The energy loss increased with the number of minor veins crossed by the signal, and approached 100% at the apex (Fig. 5). The loss of energy through several minor veins was quite well predicted when the energy loss was a function of the position of the vein. The assumption of constant loss through minor veins, however, strongly underestimated the loss of energy.

Midveins as well as minor veins showed a strong decrease of diameter from petiole to apex. This varying diameter had an important role in the transmission of energy, as the average loss of energy was constant and equal to 45% (s.d.=18%, $n=95$) for a diameter inferior or equal to 0, 2 mm (Fig. 6) and increased up to 90% for larger diameters. The loss was nearly complete when the signal crossed the midvein near the petiole.

B. Frequency-dependent losses through structural elements of a leaf

After the decomposition of each signal into five frequency levels, we pooled values to obtain two frequency ranges: lower frequencies, below 1.7 kHz, and higher frequencies, between 1.7 and 25 kHz. The reason for doing so was twofold. First, the energy content in the five levels differed substantially, while the two ranges each contain roughly half the total energy. Calculating losses from frequency levels with very low energy content was imprecise due to the small energy involved (for example, 0.6% of the total energy was contained in the level corresponding to the highest frequencies) and the high signal-to-signal variability at these levels. Second, the splitting frequency of 1.7 kHz corresponded roughly to the borderline frequency (2.5 kHz) of Meyhöfer *et al.* (1994) between ''low-frequency'' signals produced by prey moving or feeding and ''high-frequency'' signals of prey wriggling in order to escape stings of the wasps.

In a given sector, the average loss of energy was 45% (s.d.=22%, $n=42$) for high frequencies and 18% (s.d. =35%) for low frequencies. Minor veins reduced energy by 44% (s.d.=29%, $n=108$) on average for high frequencies, and by 32% (s.d.=24%) on average for low frequencies. The midvein reduced higher frequencies by 72% (s.d.=18%), and lower frequencies by 43% (s.d.=29%, $n=39$). Hence, the loss of total energy was mainly due to a loss of high frequencies. Signals changed in composition, sometimes strongly. The proportion of frequencies higher than 1.7 kHz was reduced by 18% when the signal crossed the midvein (Fig. 7). Hence, the midvein acted as a low-pass filter. The decrease was much less important for minor veins (4%) and homogeneous regions (6%, Fig. 8).

III. DISCUSSION

A. The leaf as channel in vibratory communication in prey–predator systems

The use of small dropping balls was chosen to produce short, transient vibrations similar to those observed in the leafminer system. A wavelet analysis of several signals from wriggling pupae confirmed the qualitative similarities with the signals produced by impacts, but showed quantitative differences. In particular, the energy contained in the biological signal was at least one order of magnitude less than in the signals produced by the impacts, and around two-thirds of the energy contained in biological signals was located

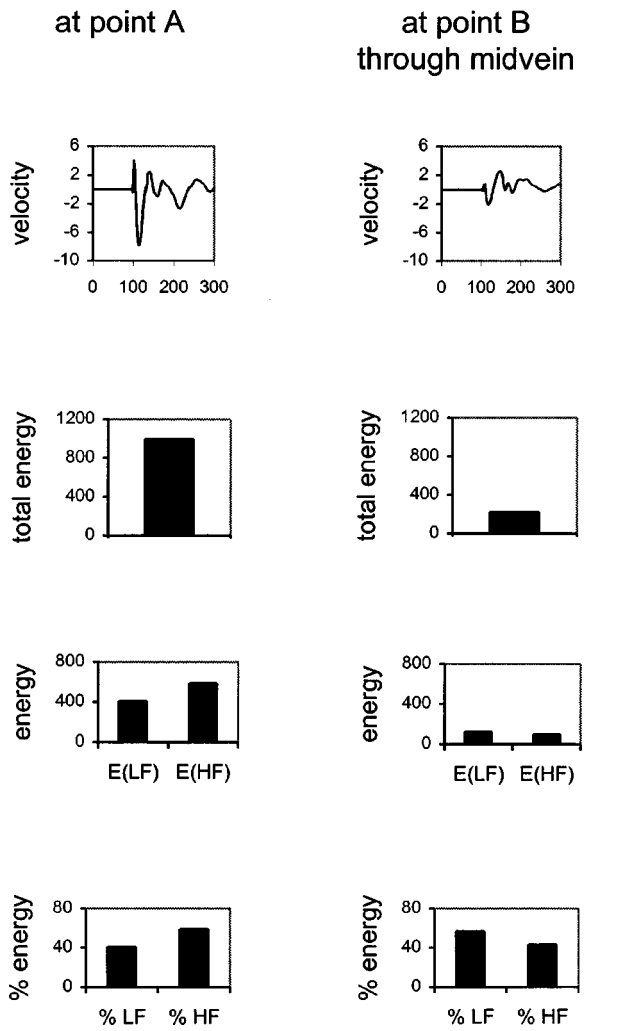


FIG. 7. Recorded signals at point *A* and point *B* and their energy analysis. The signal crosses a midvein. We represent the velocity of the signals (mm/s), the total energy, the energy contained in lower frequencies below or equal to 1.7 kHz, and in higher frequencies above 1.7 kHz, and its relative values (energy expressed in mm^2/s^2).

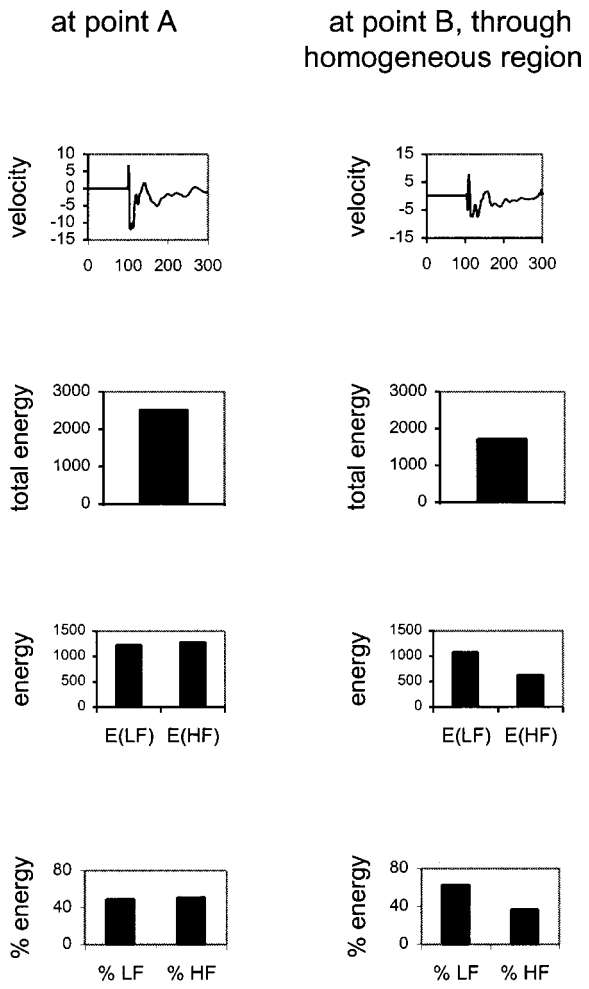


FIG. 8. Recorded signals at point *A* and point *B* and their energy analysis. The signal goes through a homogeneous region. See the legend of Fig. 7 for further information.

above 1.7 kHz. We do not understand how insects are able to create signals having a much larger proportion of high frequencies than low frequencies, a point also raised by Michelsen *et al.* (1982). A thorough study of the coupling of a vibrating insect to the plant structure may provide part of the explanation.

The strong, frequency-dependent attenuation implies that vibratory signals change over a short distance on a leaf. This applies in particular when signals cross veins, but even the leaf lamina dissipates a lot of energy. These results have two implications for understanding the leafminer system. First, they imply that the low energy signals produced when a larva moves in the mine or feeds will most likely not propagate much further than the mined area, as a mine is located between two minor veins. Hence, only specialized natural enemies may be able to perceive and react to such signals from any distance on a leaf using specifically tuned receptors. Second, a wasp on a leaf will obtain signals that vary greatly depending on its location on the leaf relative to the source. This gradient of information may be used by the wasp for locating the source. For example, the frequent and rapid swaps between several mines on the same leaf while foraging for prey may be explained by the fact that wasps receive signals coming from mines other than the one currently being searched (Casas *et al.*, 1993). In more general terms, the biological message is that the leaf architecture generates a set of options for vibratory communication, a level of complexity that would be absent in a veinless, plate-like leaf. For the sender it means stay within the same sector as the receiver if you want the signal to be well-detected, but keep veins between you and the receiver if you want to hide. For the receiver, it suggests moving to the same sector as the sender to increase signal reception.

We have shown in the previous paragraph how predators and prey can improve their situation relative to vibration reception by moving around on a leaf. Arthropods are even able to use vibration without moving. A difference of intensity of a signal between two of the legs is likely to be used by spiders living in orb webs or hunting in the vegetation to turn and locate a vibrating source (Hergenröder and Barth, 1983; Landolfa and Barth, 1996). The species *Cupiennius salei*, which hunts animals moving on the vegetation, has a leg span of 10 cm. A loss of 10 dB between two legs triggers correct turning behavior. Potentially, big arthropods can also use the frequency composition in different sensors to locate a

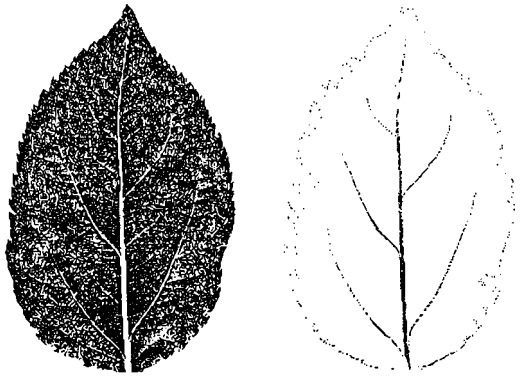


FIG. 9. An apple leaf (left) and the outline of the structures retained for modeling the propagation of waves due to impacts (right).

vibrating source, but results obtained so far are inconclusive. For *C. salei*, differences in time of arrival between legs at the peak frequency of the courtship signal are between 2 and 10 ms, which is a range known to be effective for orientation. However, prey signals contain mostly higher frequencies, leading to times of arrival between legs that are too small to trigger any behavioral response (Barth, 1998). The use of time-of-arrival and phase cues is uncertain in web spiders due to the high propagation speed of vibrations (Barth, 1998). Given the inconclusive results obtained with relatively large spiders, arthropods with a leg span of a few millimeters, such as the leafminer parasitoid, are most likely unable to use the frequency composition of vibratory signals for mate and prey recognition. Our findings suggest, however, that these arthropods could be better served using the loss of total energy, or any related statistics, which is important even over a very short distance within homogeneous regions.

B. Relative importance of structural elements in the propagation of vibrations

A real leaf is made of a few basic structural components that vary substantially at the microscopic scale. The results obtained in this study enable us to focus on the relevant level of detail for modeling the propagation of mechanical waves produced by impact-like behaviors of insects. We showed the important role of veins in determining the position-dependent energy loss through a whole leaf. We intend next to study experimentally the channeling of energy through the veins, either into other regions of the leaf or into the air. This step will enable us to characterize the radiation properties of the energy and the mechanisms by which veins act as low-pass filters. A study combining laser doppler anemometry and laser doppler vibrometry showed indeed that air movement triggered by leaf vibrations caused by such impacts could be measured up to several cm away from a leaf (Casas *et al.*, 1998).

The loss of energy in crossing a vein was also shown to be strongly dependent on the vein diameter. There is a gradient of vein diameter over a leaf, the midvein and minor veins being thicker near the petiole and becoming thinner towards the apex. Hence, we regard the diameter as being the relevant variable in determining energy loss through structures, rather than the position *per se*, and also discard the former distinction between midvein and minor veins. Veins with diameter smaller or equal to 0.2 mm dissipate as much energy as the leaf lamina. Therefore, we end up with a much simpler leaf (Fig. 9) and consider only veins of diameter >0.2 mm and the leaf lamina as being the necessary structural elements to retain in a future model of vibration propagation in apple leaves.

ACKNOWLEDGMENTS

We thank R. Meyhöfer for constructing the electrical magnet and for carrying out preliminary experiments several years ago, A. Rivero for checking the English, and the CNRS, program ''Environnement, Vie et Société,'' Grant No. 98N62/0101 for financial support.

Bacher, S., Casas, J., and Dorn, S. (**1996**). ''Parasitoid vibrations as potential releasing stimulus of evasive behavior in a leafminer,'' Physiol. Entom. **21**, 33–43.

Bacher, S., Casas, J., and Dorn, S. (**1997**). ''Substrate vibrations elicit defensive behavior in leafminer pupae,'' J. Insect Physiol. **43**, 945–952.

Barth, F. G. (**1998**). ''The vibrational sense of spiders,'' in *Comparative Hearing: Insects*, edited by R. R. Hoy, A. N. Popper, and R. R. Fay (Springer, Berlin), pp. 228–278.

Casas, J., Gurney, W. S. C., Nisbet, R., and Roux, O. (**1993**). ''A probabilistic model for the functional response of a parasitoid at the behavioral time-scale,'' J. Anim. Ecol. **62**, 194–204.

Casas, J., Bacher, S., Tautz, J., Meyhöfer, R., and Pierre, D. (**1998**). ''Leaf vibrations and air movements in a leafminer–parasitoid system,'' Biocontrol **11**, 147–153.

Cohen, L. (**1995**). *Time Frequency Analysis* (Englewood Cliffs, NJ), pp. 3–5.

Ewing, A. W. (**1989**). *Arthropods Bioacoustics*, Neurobiology and Behavior (Cambridge University Press, Cambridge).

Henry, C. S. (**1994**). ''Singing and cryptic speciation in insects,'' TREE **9**, 388–392.

Hergenröder, R., and Barth, F. (**1983**). ''Vibratory signals and spider behavior: How do the sensory inputs from the eight legs interact in orientation?'' J. Comp. Physiol. A **152**, 361–371.

Jones, H. G. (**1992**). *Plant and Microclimate*, 2nd ed. (Cambridge University Press, Cambridge), p. 428.

Landolfa, M. A., and Barth, F. G. (**1996**). ''Vibrations in the orb web of the spider Nephila clavipes: cues for discrimination and orientation,'' J. Comp. Physiol. A **176**, 493–508.

Martens, M. J. M. (**1990**). In *Physical Methods in Plants Sciences*, edited by H. F. Linskens and J. F. Jackson, pp. 1–22.

Martens, M. J. M., and Michelson, A. (**1981**). ''Absorption of acoustic energy by plant leaves,'' J. Acoust. Soc. Am. **69**, 303–306.

Meyer, Y. (**1993**). *Wavelets Algorithms & Applications* (SIAM, Philadelphia), p. 4.

Meyhöfer, R., Casas, J., and Dorn, S. (**1994**). ''Host location by a parasitoid using leafminer vibrations: Characterizing the vibrational signals produced by a leafminer host,'' Physiol. Entomol. **19**, 349–359.

Meyhöfer, R., Casas, J., and Dorn, S. (**1997**). ''Vibration mediated interactions in a host–parasitoid system,'' Proc. R. Soc. London, Ser. B **264**, 261–266.

Meyhöfer, R., and Casas, J. (**1999**). ''Vibratory stimuli in host location by parasitic wasps,'' J. Insect. Physiol. **45**, 967–971.

Michelsen, A., Flemming, F., Gogala, M., and Traue, D. (**1982**). ''Plants as transmission channels for insect vibrational songs,'' Behav. Ecol. Sociobiol. **11**, 269–281.

Misiti, M., Misiti, Y., Oppenheimer, G., and Poggi, J. M. (**1996**). Wavelet Toolbox User's Guide.

Mulkey, S. S., Kitajima, K., and Wright, S. J. (**1996**). ''Plant physiological ecology of tropical forest canopies,'' TREE **11**, 408–412.

Schuepp, P. H. (**1993**). ''Leaf boundary layers,'' New Phytol. **125**, 477–507.

Sellers, P. J., Mos, S. O., Tucker, C. J. (**1996**). ''A revised land surface parametrization (SiB2) for atmospheric GCM,'' J. Clim. **9**, 706–737.
Smith, D. M., Jarvis, P. G., and Odongo, J. C. W. (**1997**). ''Aerodynamic conductances of trees in windbreaks,'' Agric. Forestry Meteorol. **86**, 17–31.
Stewart, K. W. (**1997**). ''Vibrational communication in insects,'' American Entomologist, pp. 81–90.
Tang, S. H., Ong, P. P., and Woon, H. S. (**1986**). ''Monte Carlo simulation of sound propagation through leafy foliage using experimentally obtained leaf resonance parameters,'' J. Acoust. Soc. Am. **80**, 1740–1744.
Walh, T. J., and Bolton, J. S. (**1993**). ''The application of the Wigner distribution to the identification of structure-borne components,'' J. Sound Vib. **163**(1), 101–122.

Echolocation behavior of big brown bats, *Eptesicus fuscus*, in the field and the laboratory

Annemarie Surlykke
Center for Sound Communication, Institute of Biology, Odense University, SDU, University of Southern Denmark, DK-5230 Odense M., Denmark

Cynthia F. Moss
Department of Psychology, University of Maryland, College Park, Maryland 20912

(Received 20 March 2000; accepted for publication 14 August 2000)

Echolocation signals were recorded from big brown bats, *Eptesicus fuscus*, flying in the field and the laboratory. In open field areas the interpulse intervals (IPI) of search signals were either around 134 ms or twice that value, 270 ms. At long IPI's the signals were of long duration (14 to 18–20 ms), narrow bandwidth, and low frequency, sweeping down to a minimum frequency (F_{min}) of 22–25 kHz. At short IPI's the signals were shorter (6–13 ms), of higher frequency, and broader bandwidth. In wooded areas only short (6–11 ms) relatively broadband search signals were emitted at a higher rate (avg. IPI=122 ms) with higher F_{min} (27–30 kHz). In the laboratory the IPI was even shorter (88 ms), the duration was 3–5 ms, and the F_{min} 30–35 kHz, resembling approach phase signals of field recordings. Excluding terminal phase signals, all signals from all areas showed a negative correlation between signal duration and F_{min}, i.e., the shorter the signal, the higher was F_{min}. This correlation was reversed in the terminal phase of insect capture sequences, where F_{min} decreased with decreasing signal duration. Overall, the signals recorded in the field were longer, with longer IPI's and greater variability in bandwidth than signals recorded in the laboratory. © *2000 Acoustical Society of America.* [S0001-4966(00)03611-0]

PACS numbers: 43.80.Ka, 43.60.Cg, 43.60.Qv [WA]

I. INTRODUCTION

Pierce and Griffin, and subsequently Griffin and Galambos, conducted a series of elegant experiments around 1940 that unraveled the mystery of bat orientation in the dark (Griffin, 1958). They demonstrated that microchiropteran bats emit ultrasonic frequency sounds and listen to echoes from objects in the path of the sound beam, allowing them to orient in the dark and exploit the food sources of the night sky. Griffin (1944) called this acoustic orienting behavior ''echolocation.'' This seminal finding triggered a vast amount of research directed at understanding perception by sonar in echolocating bats. Over the past 60 years, research has emphasized the following questions: What are the characteristics of bat echolocation sounds? What adaptations are present in the bat's auditory system to analyze sonar echoes and represent the animal's environment? How does the bat's sonar receiver perform in tasks of ranging, detection, and discrimination? How does sonar signal design influence echo information processing in different bat species hunting in different environments?

Research on echolocation behavior in bats has followed two main avenues. One emphasizes laboratory experiments that introduce careful manipulation of discrete acoustic variables and measurement of behavioral performance in sonar tasks. For example, psychophysical studies determine the bat's detection or discrimination of echoes that differ in amplitude, delay, spectrum, and/or Doppler shift (for review, see Moss and Schnitzler, 1995). Other laboratory studies of echolocation behavior examine obstacle avoidance and insect capture under controlled experimental conditions (Simmons *et al.*, 1995). Neurophysiological experiments examine responses of neurons in the bat's auditory pathway to acoustic parameters relevant to biosonar (for reviews, see Suga, 1988; Popper and Fay, 1995).

The other approach to study echolocation behavior in bats involves field recordings of the signals used by different species foraging in different habitats. There are around 700 species of echolocating bats, and they differ in their sonar signal designs for foraging under different conditions. Bat sonar signals may contain constant frequency (CF) and frequency modulated (FM) components, and these signals range from less than 1 ms to more than 100 ms in duration (Simmons and Stein, 1980; Neuweiler, 1984; Fenton, 1995). Most species that have been studied to date show changes in the parameters of their sonar signals (e.g., duration, bandwidth, and repetition rate) with foraging conditions, such as their proximity to vegetation, water, and buildings (Fenton, 1986; Neuweiler, 1984; Schnitzler and Kalko, 1998). It is widely believed by researchers in the field that sonar signal design reflects the bat's active control over important acoustic information gathered from sonar echoes (Griffin, 1958). Thus one can gain insights to the basics of echolocation by comparing sonar signals produced by different bat species for echolocation under similar and different environmental conditions (Neuweiler, 1984).

Only a small number of the extant bat species have been the subject of extensive field studies. Among the insectivorous species studied, there are some general patterns in acoustic behavior that accompany the pursuit and capture of prey (Simmons *et al.*, 1979). In particular, insectivorous bats produce sonar signals that decrease in duration and increase

in repetition rate as they approach a prey item. The bat's acoustic behavior during insect capture is typically divided into the search, approach, and terminal phases, largely distinguished by the repetition rate of the bat's sonar signals. During the search phase, sonar signals are typically produced at a rate of 2–10 per sec. The repetition rate may increase to about 80 sounds per sec during the approach phase and to as high as 200 sounds per sec during the terminal phase. The terminal phase signals occur at such a high rate that it is often referred to as the terminal buzz (Griffin, 1958).

A common goal of laboratory and field studies of echolocation behavior in bats is to characterize the acoustic signals used to probe the environment and to understand perception by sonar (Simmons *et al.*, 1975). While there are similarities in some broad characteristics of sonar signals produced by different insectivorous species, there are also considerable differences in time-frequency structure, bandwidth, duration, and repetition rate of sonar signals recorded from different species under different foraging conditions. Furthermore, behavioral data on echolocation task performance reveal species-specific differences that appear related to sonar signal features. Thus one cannot simply take information about the natural acoustic behavior of one bat species to interpret the results of laboratory experiments on another bat species. Therefore, it is most striking that almost no field recordings have been published from the big brown bat, *Eptesicus fuscus*, one of the most widely studied bat species in laboratory experiments (Covey and Casseday, 1995; Simmons *et al.*, 1995; Simmons, 1973; Moss and Schnitzler, 1995). Apart from Griffin's thorough, but old, data (Griffin, 1958), the published information on acoustic behavior of *E. fuscus* in the field is limited to bat detector recordings (Betts, 1998, or web sites, e.g., http://talpa.unm.edu/batcall) and cursory data in reviews and papers focusing on other aspects of echolocation (Obrist, 1995; Simmons *et al.*, 1979; Simmons, 1987). The detailed acoustical data on sonar signals produced by *E. fuscus* come from recordings of bats in the lab (Hartley, 1992). However, it has been demonstrated repeatedly that bats do not emit the same signals in the confined environment of a laboratory as they do in their natural habitat (Griffin, 1958).

Thus one purpose of the present study was to provide more detailed reference data on the natural acoustic behavior of *Eptesicus fuscus* in the wild. A second purpose was to compare this species' acoustic behavior in the wild with the echolocation signals it produces while catching tethered insects in a large laboratory flight room.

II. METHODS

A. Field locations

We recorded echolocation sounds from *E. fuscus* at three different sites in Maryland, in areas where we identified *E. fuscus* colonies and/or foraging behavior.

At the first field site, A (Rockville, North of Washington, D.C., recording date: Aug. 3, 1999), we recorded cruising *E. fuscus* as they flew out of their roost on the way to their hunting grounds. The colony roosted in the attic of a townhouse on the shore of an artificial lake. Leaving the roost, the bats flew out from a small opening under the roof, out over the lake, and turned left to fly past the recording microphone (see below) before disappearing over some trees on the opposite brink. Hence, in general our recordings were made with the microphone at $-90°$ with respect to the general flight path of the bats. We started the recordings right after sunset (8:45 pm) when the first bats flew out and stopped at 9:50 pm, before the bats returned. Thus it is highly unlikely that we recorded any single bat more than once, and the risk of pseudoreplication is very small.

The second site, B (Gambrill, 10 km NW of Annapolis, Maryland, recording date: Aug. 8, 1999) was an open hunting area. There was a large grass lawn stretching away from a house. The lawn was lined by large old trees, and an open field extended beyond. The bats flew and hunted at different heights at this site, from 4–5 m above the ground to well above the level of the tree tops (>10 m), and our recordings included signals from several bats occupying the area at the same time.

The third site, C (Greenbelt, close to University of Maryland, College Park, recording date: Sept. 14, 1999) was a more confined hunting area, a road (7-m wide), running through a wooded area. We identified this site as ''cluttered,'' in accordance with Kalko and Schnitzler, 1993. Along the road were street lamps, and the bats hunted primarily around the lamps, ~8 m above the ground. Here, as at site B, we have simultaneous recordings of several bats. Thus we know we recorded more than one bat, but we have no control over which bats we recorded, and some of the sequences we chose for analysis might well be from the same bat.

B. Lab recordings

The lab recordings were carried out under long wavelength lighting (>650 nm, red filters, Reed Plastics, Rockville, MD) in a $6.4\times7.3\times2.5$ m carpeted flight room, with walls and ceiling lined with acoustic foam (Sonex 1, Illbruck). Two *E. fuscus* were trained to capture tethered mealworms. The worms hung from the ceiling of the flight room by a very thin thread. The position and movement of the worm varied from trial to trial. The bats' flight and capture behavior were recorded on two gen-locked high-speed video recorders (Kodak MotionCorders) at 240 frames/sec. The bat's three-dimensional (3-D) flight path was later reconstructed using commercial motion analysis software (Motus, Peak Performance Technologies).

C. Sound recordings

We used two different recording setups. At sites A and B we used an Ultrasound Advice SM2 microphone and SP2 amplifier (sensitivity is flat up to 40 kHz and decreases gradually by 5 dB from 40 to 100 kHz). The signals were bandpass filtered from 10–100 kHz (filter slope -110 dB at $1.5f_c$) and further amplified using a Stanford Research Systems Model SR 650 digital filter. At site C we used a battery operated system consisting of a 1/4″ G.R.A.S. microphone model 40 BF, amplified (40 dB) and high-pass filtered (13 kHz) through a G.R.A.S. 12AA microphone power supply.

This system is linear (±1 dB from 15–100 kHz). In all cases the microphone was mounted on the end of a ~2-m-long thin rod. The signals recorded by the microphone were continuously A/D converted on-line (12-bit, sampling rate 250 kHz) using a battery operated IoTech Wavebook 512. The signals were stored in a ring-buffer (FirstInFirstOut, FIFO). The Wavebook had 128 Mbytes of random access memory (RAM) and was usually set at 6 sec pretriggering time and 1 sec post-triggering time. The Wavebook was controlled by a laptop computer (AST Ascentia P series or IBM Thinkpad 600). Thus when we picked up strong signals on a bat detector (D240 Pettersson Electronics), we began signal acquisition with the Wavebook, storing the 6 sec preceding and the 1 sec following trigger. At all recordings sites, the data storage system was battery operated. Hence, at site C the entire recording chain, including filtering and amplification of the microphone output, was battery operated, resulting in a lower noise floor than at sites A and B, where we used AC to power the Stanford Research Systems filter/amplifier.

In the lab, a similar system was used. We used two microphones, separated by approximately 3 m, each positioned about 30 cm above the floor. One microphone was supplied by Ultrasound Advice (as at sites A and B) and the other by ACO Pacific (1/4″ microphone, model 7016PM). The signals were bandpass filtered (10–100 kHz) and amplified, using a Stewart Electronics Model VBF 44 filter (−110 dB at $1.5f_c$). The signals were digitized directly on two channels of an IoTech 512 Wavebook (sample rate 240 kHz per channel, over an 8.2-sec period, 7.2-sec pretrigger, and 1-sec post-trigger), and controlled by a Dell Inspiron laptop computer. The channel containing recordings with the best signal-to-noise ratio in the terminal buzz phase of the capture behavior was later chosen for analysis.

D. Sound analysis

We analyzed the sounds using a commercial software program: BATSOUND (Pettersson Electronics). The signals were displayed simultaneously as spectrograms and oscillograms (time traces). The interpulse intervals (IPI) and signal durations (dur) were measured from the oscillograms, while using the spectrograms to distinguish between signals and their echoes. The spectrograms were made of consecutive fast Fourier transforms (FFT's) with a 90% overlap. Usually, a 256-point FFT was chosen as a compromise between good frequency resolution (977 Hz) for the shallow sweep tail of the signals and good time resolution for the steeper FM sweep of short signals and of the initial part of longer signals. For inspection of finer details in either time or frequency, the FFT window was adjusted accordingly.

In order to follow frequency changes correlated with habitat and behavior, we chose to measure the minimum frequency ($F_{\min}$) of all signals, since low frequencies are less subject to excess attenuation than high frequencies. In most cases $F_{\min}$ was measured from the spectrograms as the lowest frequency that clearly had more energy than the background noise. This method was repeatedly checked by producing spectra and measuring $F_{\min}$ from those as the low-frequency limit of $BW_{-15\,\mathrm{dB}}$, the bandwidth of the spectrum at −15 dB relative to the intensity of the peak frequency.

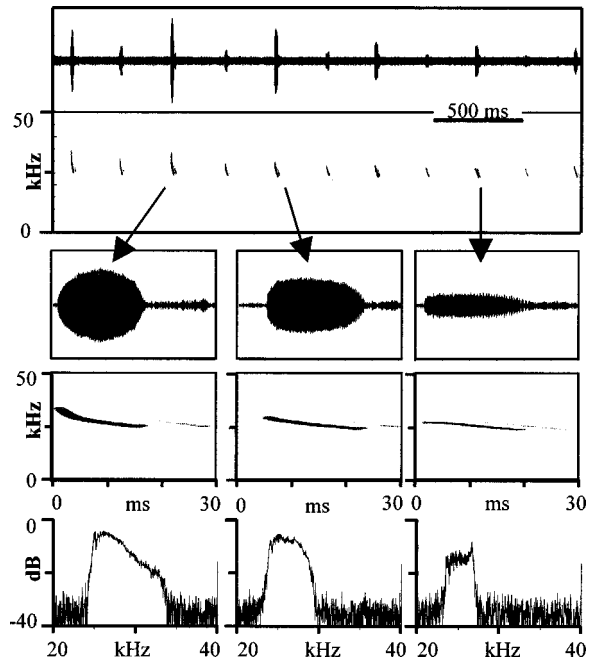


FIG. 1. Signals recorded from cruising *E. fuscus* leaving their roost at area A. The amplitudes alternate. The interpulse intervals (IPI's) are around 270 ms. The signal duration increases through this recording. The spectrograms and spectra of the three signals shown in the zoom below show decreasing frequencies (both Fmin and Fmax) and bandwidth with increasing duration.

Signal bandwidths are often given as the width of the full spectrum above the noise. However, such a bandwidth measure will depend upon the recorded signal level, i.e., the more intense the signal (the closer the bat) the broader the bandwidth. We chose a standard −15-dB bandwidth to try to use a bandwidth measure that is independent of differences in signal levels. For search and approach signals the $F_{\min}$ measured from the spectrogram corresponded very well with the $F_{\min}$ measured from the spectrum of the same signal: the discrepancy was less than 1 kHz. However, at the lowest signal levels (terminal buzz signals recorded from bats at a distance) $F_{\min}$ was often difficult to measure. The variations could be up to 2–3 kHz, and the values should be taken with caution. To examine the relation between the changes in $F_{\min}$ with changes in peak frequency and maximum frequency (and thus in overall bandwidth), we produced spectra of sounds from the best field recordings and from the lab recordings. For those spectra we used a rectangular FFT window that was adapted to the signal duration, e.g., 8192 points (32.8 ms at 250-kHz clock rate) for the longest signals and 512 points (2.13 ms at 240-kHz clock rate) for the shortest signals.

III. RESULTS

A. Cruising bats

When the bats flew out of the roost (site A), they produced long signals. The amplitudes often alternated between high and low, which might be due to the bats scanning up and down or from side to side and changing the direction of their sound beam with respect to the microphone (Fig. 1). For analysis we selected six recordings with only one bat and a good signal-to-noise ratio. The distribution of interpulse

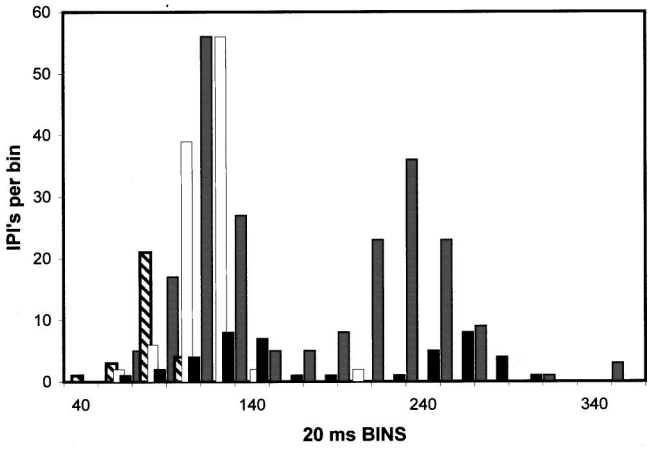


FIG. 2. Interpulse interval histogram. Bin width: 20 ms. The labels give the lower limit of each bin, e.g., 40 is the bin from 40 to 60 ms. The distribution is bimodal for the cruising bats (black bars) and the bats hunting in the open area, B (gray bars), but shows only a single peak for the bats hunting in the cluttered area, C (white bars) and in the lab (diagonally striped bars).

intervals (IPI's) was bimodal (Fig. 2) with a peak at 120–160 ms (avg: 134 ± 27 ms, $N = 24$, where N is the number of intervals) and another peak at approximately twice that value (avg: 270 ± 19 ms, $N = 19$). To calculate these averages we arbitrarily divided the IPI's into two populations, above or below 200 ms, based on the breakpoint in the histogram (Fig. 2). The signal durations changed with the IPI's. At long IPI's the signal durations were between 12 and 19 ms, whereas at IPI's around 120–160 ms, the durations were only between 6 and 13 ms (Fig. 3). The $F_{\min}$ is also correlated with IPI and thus with signal duration, such that when signal duration increase, $F_{\min}$ decreases (Figs. 1 and 3). The correlation between signal duration and $F_{\min}$ is strong [Fig. 3(b)], but does not necessarily imply dependence of one parameter on the other. The bats maintained control over the frequency and could change the frequency without changing the duration (see Fig. 4).

The spectra and spectrograms in Fig. 1 reveal some typical characteristics of the longest signals. The three analyzed signals have durations of 16, 18, and 19 ms, respectively. With increasing signal duration both the $F_{\min}$ and the bandwidth decrease. The initial, steeper sweep rate in the shortest signal is nearly absent in the longer signals. Accordingly, the spectrum changes from having more energy at low frequencies in the shorter signals to a more even distribution of energy across the spectrum in longer signals. Finally, the longest signal starts with an almost constant frequency component, resulting in a small peak at the high-frequency end of the (very narrow) spectrum. However, again the recordings demonstrate how flexible the bats are. For example, comparison with Fig. 5 shows that even the longest signals may start with a short downward sweep.

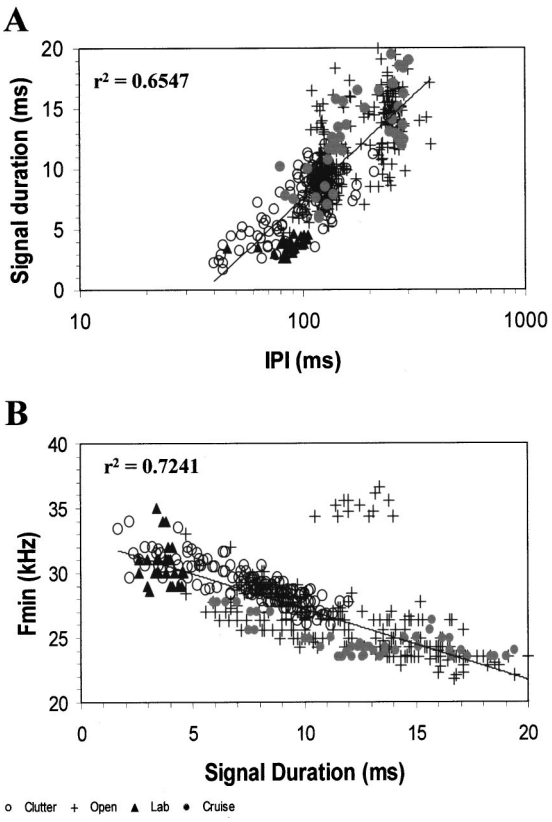


FIG. 3. (a) Correlation of signal duration and IPI. The r^2 of the log-regression line is 0.65 ($y = 7.44 \ln(x) - 26.73$). (b) Negative correlation between signal duration and $F_{\min}$, for all nonbuzz signals. The linear regression line ($r^2 = 0.72, y = -0.55x + 32.66$) is calculated without the 13 high-frequency signals far above the line recorded from a single bat in the open area. Open circles: C, the cluttered, wooded area. Crosses: B, the open area. Closed, gray circles: A, cruising bats. Black triangles: Lab recordings.

Since the water reflected strong echoes at site A, we could get a very rough estimate of the bats' flight elevation by measuring the echo delay, since all bats followed more or less the same route over the water. Our microphone was approximately 3 m over the water surface. Thus for example, the delay of the reflected sound from a bat at a horizontal distance of 15 m would be ~3 ms at a flight elevation of 3 m and 10 ms at a flight elevation of 10 m. We could measure

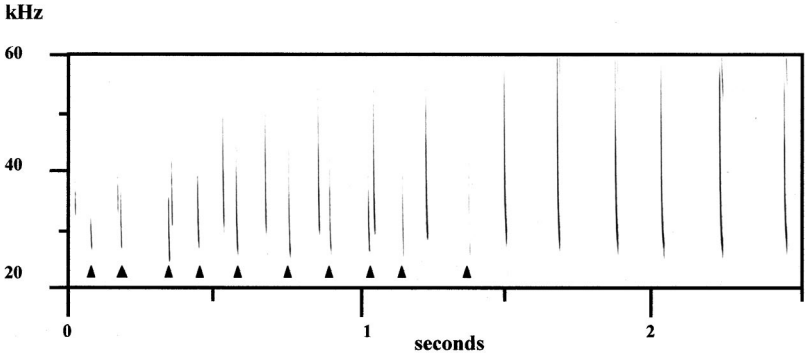


FIG. 4. A simultaneous recording of two bats. At first, their $F_{\min}$'s are separated by 6–8 kHz. Around 1 sec the recorded intensity of the low-frequency bats drops and that of the high-frequency bat increases, while it starts lowering its $F_{\min}$.

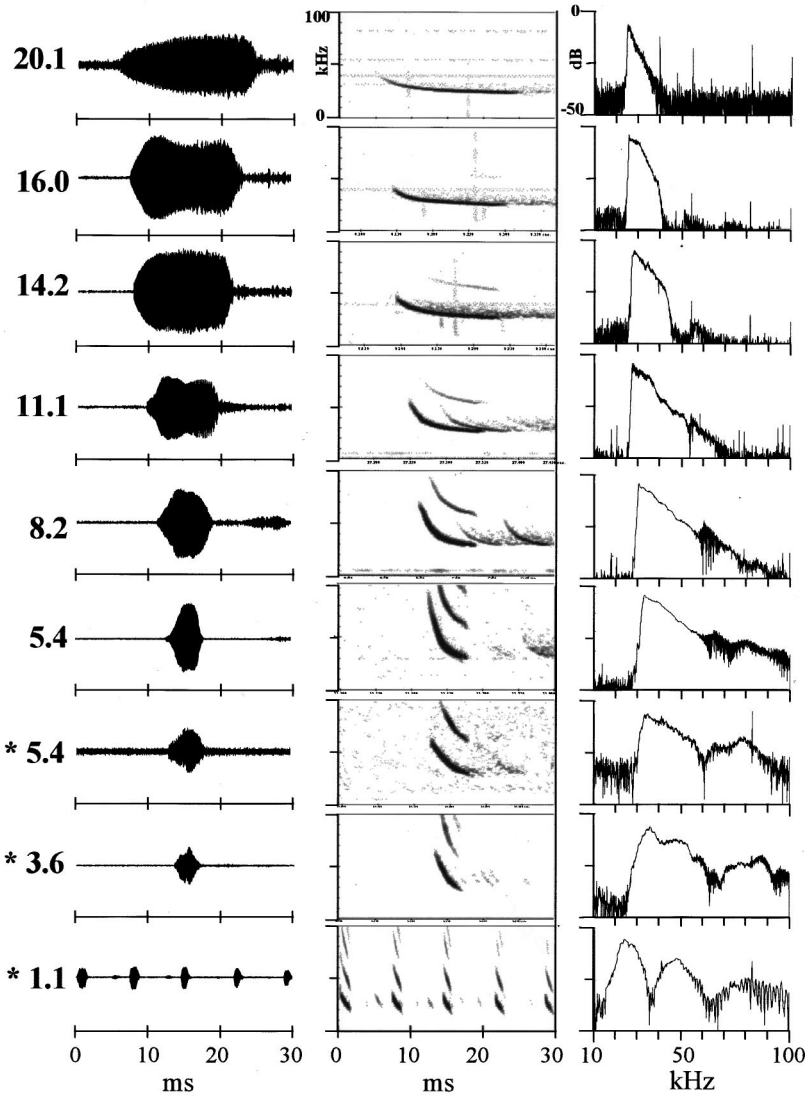


FIG. 5. Typical sounds. Oscillograms, spectrograms, and spectra of sounds with durations from 20.1 to 1.1 ms. Durations in ms are given at the left. Asterisks mark lab recordings. Oscillograms and spectrograms show 30 ms. All spectrograms go from 0 to 100 kHz. All spectra go from 10 to 100 kHz, and show relative dB on a scale from 0 to −50 dB, with the spectra adjusted to a peak value of ~−5 dB. The steep slope on the high-frequency side of the spectrum (especially 16.0 and 14.2 ms) indicate that the lack of high frequencies is a feature of the signals, not a consequence of excess atmospheric attenuation of higher frequencies. In some signals echoes from surroundings are prominent (e.g., 11.1 and 8.2 ms).

the delay directly from the spectrograms (Fig. 1) or indirectly from the interference pattern between direct and reflected sound. The interference creates typical amplitude modulations of the time signals and characteristic notches in the spectra separated by (1/delay ms) kHz, e.g., 3.3 kHz at 3-ms delay between direct and reflected sound (Kalko and Schnitzler, 1989a). Echo delays ranged between 3 and 15 ms and the longer IPI's and longer signals were produced by bats flying at relatively high elevations. In six analyzed recordings only one bat produced short signals (6–8 ms), and this bat was also the only one flying low over the water, indicated by echo delays of 3–3.5 ms. The other five flew higher (echo delays 8–15 ms) and produced longer signals (11–19 ms).

B. Bats hunting in the field

We recorded insect pursuit sequences at sites B and C (Figs. 6 and 7). The signals in the search phase were different in the two habitats. At site B, the open site, the signals (13 sequences analyzed) resembled those of the cruising bats from site A, with IPI's of either around 130–150 ms (avg: 134 ± 15 ms, $N = 100$), twice as long (249 ± 24 ms, $N = 101$), and in a few cases, around three times the basic IPI ($N = 4$: 336,363,376,378 ms) (Fig. 2). At site C, the more cluttered, wooded environment, all the bats' IPI's were short, clustering around 120 ms (avg: 122 ± 19 ms, N 120, 10 sequences analyzed) (Fig. 2). Thus the average IPI at C was shorter than the average of the short IPI's at B ($p \ll 0.01$, two-tailed t-test for equal variances).

IPI, signal duration (dur), and $F_{\min}$ of the search signals recorded at sites B and C were correlated in the same way as measures from signals recorded from the cruising bats at site A. Thus pooling all signals confirmed that signal duration decreases with decreasing IPI. The log regression line [Fig. 3(a)] for the correlation between IPI and dur had a higher $r^2(0.65)$ than the linear regression line ($r^2 = 0.61$), and can be explained by the observation that the increase in signal duration levels off at very high IPI's. The reverse correlation between duration and $F_{\min}$ was also seen for the hunting bats at B and C as well as for the cruising bats at site A [Fig. 3(b)]. Only a single recording gave $F_{\min}$ values that fell outside the main cluster: one (''soprano'') bat recorded at the open site, B, emitted 10–12-ms signals with IPI's of 120–140 ms, thus showing the same correlation between IPI and signal duration, but with unusually high $F_{\min}$, around 35–36 kHz.

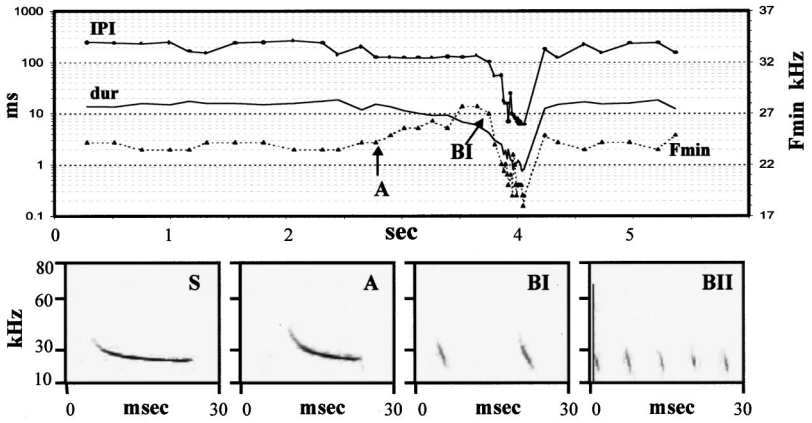


FIG. 6. Pursuit sequence from B, the open area. The upper panel shows the IPI and signal duration (dur) in ms (left *y*-axis, log-scale) and F_{min} (dotted line, triangles) in kHz (right *y*-axis, lin-scale). Below are shown spectrograms of a search signal (dur=18.5 ms), an approach signal (dur=13 ms), two Buzz I (dur=2.4 ms), and five Buzz II (dur=1 ms) signals. Search phase signals had IPI's of ~240 ms and F_{min} around 23 kHz. Note that in this case IPI is kept relatively constant around 120 ms throughout the approach phase.

The start of the approach phase was defined as the first signal in a sequence with monotonic changes in the parameters IPI, dur, and F_{min}, indicating that the bat had reacted to a potential prey (Schnitzler *et al.*, 1987). The dur and IPI decreased and F_{min} increased gradually in all recorded approach phases. The IPI and signal duration in the beginning of the approach phase were highly variable, depending on the signal parameters recorded in the preceding search phase. Therefore it was impossible to define a specific value of any of these parameters that would unequivocally indicate the start of the approach phase. Thus the approach phase can only be defined relative to the preceding search phase of the same sequence. However, the end of the approach phase was more consistent across all recorded sequences, with IPI's around 50 ms (ranging from 43 to 66 ms) and signal durations ranging from 5.5 to 2.9 ms (avg. 3.9 ms, S.DEV.=0.8 ms, $n=9$ sequences, see Fig. 8).

Over the course of the approach F_{min} increased by ~3–4 kHz. The beginning of the buzz was characterized by further decreases in dur and IPI. However, the most characteristic feature of the buzz was the reversal of the correlation between signal duration and F_{min}; at the buzz onset, F_{min} started decreasing with decreasing signal duration, thus allowing for an exact definition of the start of Buzz I based on F_{min}. The start of the Buzz II phase was characterized by an abrupt decrease in IPI. Over just 1–2 cries the IPI reached a minimum of 6 ms (pulse repetition rate of 167 Hz) and remained constant throughout the entire duration of Buzz II. F_{min} also decreased rapidly at the transition from Buzz I to Buzz II [although not as sharply as in, e.g., *Pipistrellus kuhli*

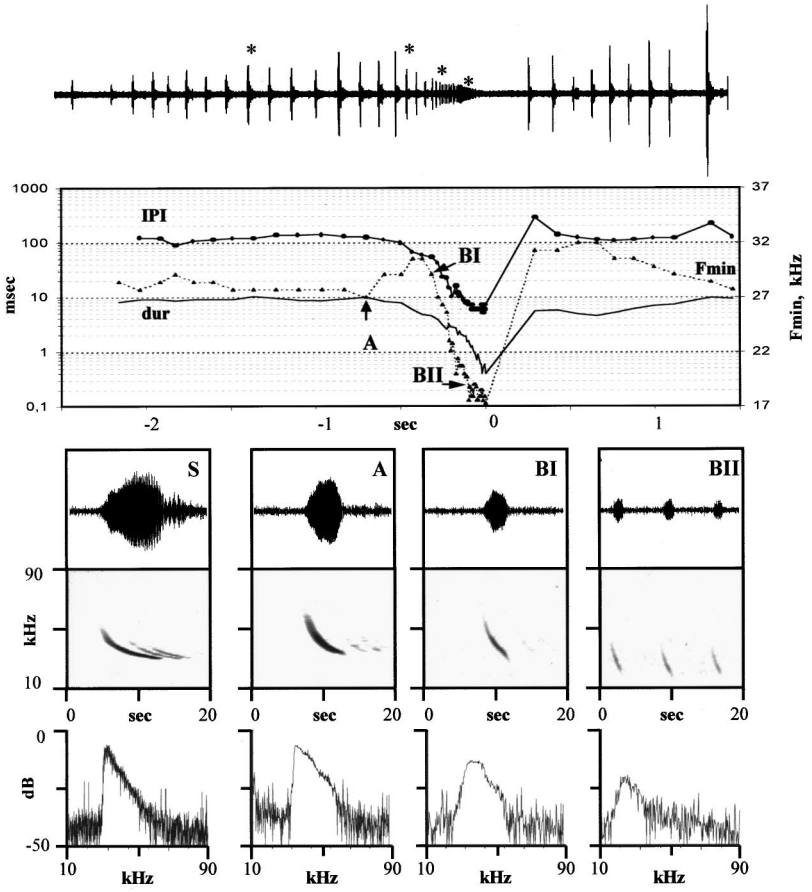


FIG. 7. Pursuit sequence from C, the cluttered area. The upper panel, an oscillogram, clearly shows how the amplitude drops in the buzz. The panel below shows IPI, dur, and F_{min} as in Fig. 6. Below are oscillograms, spectrograms, and spectra of the search, approach, Buzz I, and Buzz II signals marked with asterisks in the oscillogram.

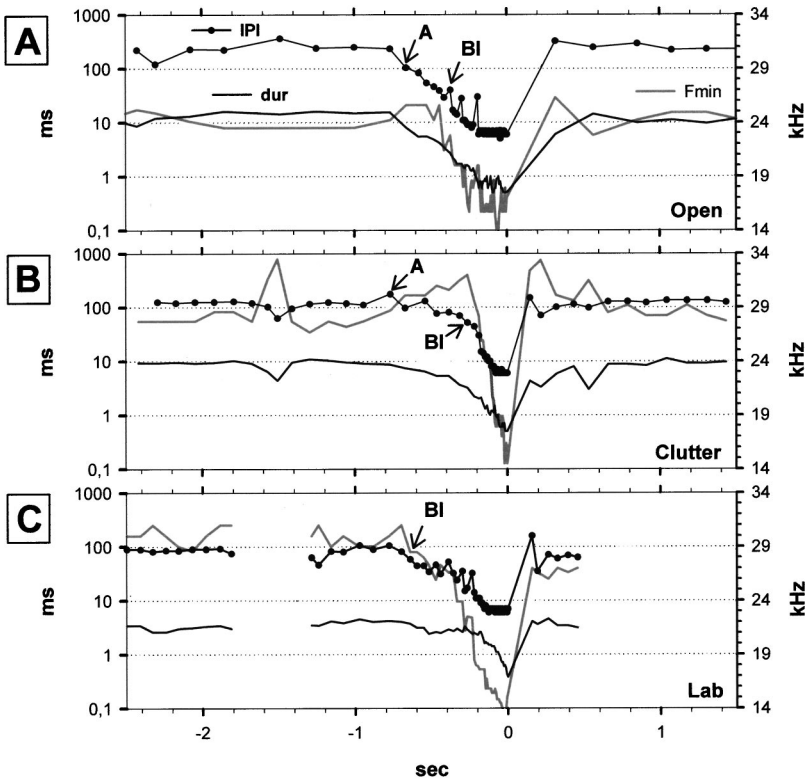


FIG. 8. Pursuit sequences from (a) the open area, (b) the cluttered, wooded area, and (c) the lab. The sequences in (a) and (b) are not the same as in Figs. 6 and 7. The IPI and dur are in ms (left *y*-axis, log-scale). The $F_{\min}$ (gray line) is in kHz (right *y*-axis, lin-scale). The start of approach and Buzz I is marked by arrows and ''A'' and ''BI.'' From upper to lower panel, the signals well before the buzz show a gradual decrease in IPI and dur and an increase in $F_{\min}$. In contrast, the signals in the final Buzz-phase are very similar in all three situations.

(Schnitzler *et al.*, 1987) or *Myotis siligorensis* (Surlykke *et al.*, 1993)]. $F_{\min}$ continued decreasing during Buzz II, but at a slower rate, to end at minimum values of 14–17 kHz (Figs. 6–8). However, there was no sudden change in signal duration at the transition from Buzz I to Buzz II. Instead, signal duration decreased gradually across Buzz phases I and II, starting in Buzz II around 1.2 ms and ending at 0.6–0.8 ms.

The bats' echolocation behavior was flexible, and we recorded variations in all of the above-mentioned parameters. For example, Fig. 6 shows an initial decrease in the bat's IPI from around 260 to 120 ms, and then the IPI remained around 120 ms throughout the approach phase, while $F_{\min}$ rose and signal duration decreased. The duration of the approach and buzz phases varied. The approach phases we recorded lasted between 320 and 620 ms. The buzz phases (Buzz I+Buzz II) lasted between 170 and 660 ms. Also, in other respects the bats demonstrated a great deal of vocal control: We often observed breaks in the buzzes [see, e.g., Figs. 6, 8(a)] and a simultaneous increase in $F_{\min}$, indicating that the breaks were not due to drop-outs in the recordings.

Our recording equipment did not allow us to determine the distance to the bats, and thus we could not determine the emitted intensity. However, the relative amplitude of signals recorded from *E. fuscus* in the present study can be compared within each sound file (6 to 9 sec of continuous recording). The nine pursuit sequences we analyzed show relatively high amplitudes, both before and after the buzz phase, thus suggesting that changes in relative amplitude are not due to changes in distance between the bat and the microphone. These data often show alternating signal amplitudes associated with the search phase similar to the alternating amplitudes for the cruising bats at site A. The recordings also reveal a clear reduction of emitted signal intensity through the terminal buzz phase. In the nine completed pursuit sequences we recorded from sites B and C, the average amplitude in Buzz I was −16 dB and in Buzz II −19 dB relative to the average amplitude of search and approach signals just before and after the terminal buzz (Fig. 7).

C. Bats flying in groups

None of the recordings with vocalizations of more than one bat was included in the data presented above, but we analyzed seven recordings of more than one bat, three from site A, two from B, and two from C. In most cases, the individual bats can be distinguished by a combination of differences in call characteristics, delay of echoes (e.g., from the water surface), and temporal patterning (IPI). The IPI's of bats flying in close proximity were less regular, and both IPI's and signal durations were shorter than those recorded from bats flying alone. Two of the recordings showed overlap in $F_{\min}$ between individual bats. However, in both of these recordings, one of the bats started pursuing an insect, shortening the signals and modifying the $F_{\min}$ up (through approach) and down (through the buzz).

The five remaining recordings of more than one bat only included search/cruise signals. These simultaneous recordings of bats flying in close proximity revealed individual calls that differed in $F_{\min}$ by 2–10 kHz, which was easily distinguished by the human ear in a slowed-down playback. Figure 4 shows an example recorded at site B, where the separation between the $F_{\min}$ of two bats is 6–8 kHz in the beginning of the recording. Here both bats emit short signals (4–7 ms), thus in the case of the low-frequency bat, demonstrating that the bats have sufficient control of all signal pa-

rameters to be able to produce signals with low F_{min} even at short durations. Around 1 sec into the recording, the bats start separating, indicated by the decreasing signal level of the low-frequency bat and the increasing level of the high-frequency bat. As the low-frequency individual disappeared, the former high-frequency bat gradually lowered the F_{min} down to values comparable to those emitted by the other bat in the beginning of the recording.

D. Insect capture in the lab

The sonar signal sequences recorded from bats flying in the lab differed from those in the field, especially in the absence of a true search phase [Fig. 8(c)]. The bats never produced the long narrow-band signals with low repetition rate that characterized search phase signals in the field. Even long before moving in on the prey (the tethered mealworm), the IPI's [avg: 88 ± 12 ms, N (number of IPI's) $= 29$ from two recordings, one from each bat] corresponded to those recorded from bats in the approach phase in the field. The signals were generally shorter (2–4 ms) than those at similar IPI's in the field [Fig. 3(a)].

In the lab, measures of F_{min} corresponded well with F_{min} for signals of similar duration recorded in the field [Fig. 3(b)]. We never observed the increase in F_{min} that accompanied the start of the approach phase in the field, probably because the bats were in constant approach phase. Thus it was difficult to define the approach phase at all using the same criteria as in the field recordings.

The start of the terminal phase was characterized by decreases in signal duration, F_{min}, and IPI, as in the field. There was large variation in the duration of the terminal phase recorded in the lab insect capture sequences, but most often the terminal phase was long compared to field recordings. We analyzed buzzes from 11 recordings of the two bats (5 and 6, respectively). The Buzz (Buzz I + Buzz II) ranged from short (230 ms) up to as much as 1.05 sec. In contrast, the other signal parameters of the lab-buzzes, IPI, duration and F_{min}, resembled those from the field recordings. The good signal-to-noise (S/N) in the lab recordings allowed us to determine the signal parameters at the end of the buzz phase more accurately than in the field recordings. The bandwidth of the signals was broader, with more energy at high frequencies, compared to signals of similar duration from the field recordings. This is probably not due to frequency differences in the emitted signals, but a consequence of the short distance between the bat and microphone.

We used the high-speed video recordings to determine the wing beat rates. The average wing beat period was 83 ms (12 Hz) (S.D. $= 17$ ms, $N = 98$ wing beat periods, from ten video recordings of several bats, not including the two bats above).

E. General signal features

The relations between IPI, duration, and F_{min} were consistent in all the recordings. Thus search signals and cruising signals from all three field locations and the sonar signals (well before the capture) from the lab recordings were pooled to examine the relations between the signal parameters of duration, IPI and F_{min} (Fig. 3). There were clear habitat-dependent differences in distribution of, e.g., IPI and thus signal duration and F_{min}. However, signals of equal duration from the different recording sites were quite similar. Thus in spite of individual and situation-dependent variations, any given signal duration corresponds to a typical signal design with a predictable F_{min} and bandwidth (Fig. 5) for bats flying alone. All recordings showed that the shorter the signal duration and IPI, the broader the signal bandwidth.

IV. DISCUSSION

A. General results

The present study generated three main findings. One, we confirmed preliminary reports that *E. fuscus*, searching for prey in the field, produce signals that are never recorded in the lab. Signal durations of up to 18–20 ms with narrow bandwidth and low minimum frequency are not rare exceptions, but routinely produced in open areas at IPI's of 260 ms or more. Second, we identified a strong correlation between IPI, signal duration, and frequency of search signals. Finally, we found that bats catching tethered insects in the lab do not produce search signals, but only approach and terminal buzz signals. The differences between lab and field recordings we report here are consistent with Griffin's early data and predictions. Modern, more sensitive sound recordings have also enabled us to extend Griffin's work by detailing some characteristics of the bat's acoustic behavior that have not been reported previously (see below).

B. Pulse repetition and wing beat rate

We found statistically reliable correlations among IPI and signal duration and F_{min}. The reliable correlations among these parameters suggest that long IPI measures are not due to missing signals in the recordings. Griffin (1958) described similar long IPI's. Furthermore, Schnitzler *et al.* (1987) and Kalko (1994) also described a similar bimodal distribution of IPI's for pipistrelle bats. Kalko ascribed the long IPI's to the skipping of signals at every second wing beat, or the execution of short gliding flights, but with signals still emitted in phase with the wing strokes. In *E. fuscus* the fundamental periodicity of the IPI histogram is 120–140 ms. This is likely to correspond to the wing beat rate of *E. fuscus* flying in the field. It is energetically most favorable for a bat to emit its signals in phase with the down strokes of the wings (Speakman and Racey, 1991). Our wing beat rates from the lab are in accordance with another more thorough study in the same flight room (Wilson and Moss, 2000). The data suggest that bats flying in the laboratory maintain the phase relation between wing stroke and pulse emission. Hence, the difference between lab and field in bats' sonar signal repetition rate indicates that also the wing beat frequency is higher in the laboratory than in the field.

C. Signal duration and frequency

The maximum signal duration we recorded was 20 ms, but 17–19 ms seems to be the more typical upper limit for duration, even when the IPI was increased to 360 ms. The

fairly long distance to the bats producing the longest signals probably contributes to the absence of a second harmonic, but does not explain the very steep high-frequency roll-off in the spectra (Figs. 1, 5). Thus the narrow bandwidth of the first harmonic is a signal feature. This confirms the data on the European bat, *Eptesicus serotinus*, by Jensen and Miller (1999). They used a microphone mounted on a 15-m-high mast, to show that high-flying bats actually produce long low-frequency signals with narrow bandwidth. The lack of high frequencies in Jensen and Miller's data cannot be attributed to low-pass filtering of the signals by excess attenuation, because their microphone was positioned at an elevation near the high-flying bats. In general, there seems to be a strong similarity between signal characteristics of *E. fuscus* and those reported for the European *E. serotinus. E. serotinus* emits source levels as high as 125 dB SPL [root-mean-square (rms)] at 10 cm (Jensen and Miller, 1999). It would be of interest to compare these signal levels with those of *E. fuscus*. However, none of the methods we employed allowed us to determine the source levels of the emitted signals, and we propose that these measures should be taken in future field studies of *E. fuscus* echolocation cries, ideally combined with photographic documentation of their capture behavior.

There was a statistically reliable correlation between signal duration and $F_{\min}$. Doppler shifts of $F_{\min}$ due to the bat's velocity relative to the microphone are likely to have contributed to some variation of the data. However, even flight speeds of 10 m/s would not change the frequency by more than around 0.7 kHz (at 25 kHz) and the changes in $F_{\min}$ reported here exceeded that value considerably and were correlated with changes in signal duration. The maximum frequency of the signals, $F_{\max}$, decreased when duration and IPI increased. With increasing signal duration, $F_{\max}$ decreased more than $F_{\min}$, producing a decrease in overall signal bandwidth (see Figs. 1 and 5). However, values for $F_{\max}$ in particular should be interpreted with caution, because we had no way of measuring the distance to the bats. Since atmospheric attenuation is more severe at the higher frequencies of the FM sweep produced by *E. fuscus* (Lawrence and Simmons, 1982), it is likely that many of the cries we recorded at long distances contained more energy at high frequencies, in particular more energy in the second harmonic, than our spectra showed. Since the longest signals were emitted by bats flying in the open far away from vegetation and ground, and thus from the microphone, the lack of second harmonic is more pronounced in long signals than in short (compare, e.g., 20.1 and 16 ms with 8.2 and 5.4 ms in Fig. 5). This might also have affected measures of signal duration by filtering out the beginning of the signals. The directionality of *E. fuscus* signals increases with frequency (Hartley and Suthers, 1989), thus also contributing to a high-frequency loss, since in most cases the bats would not consistently direct their vocalizations toward the microphone.

Kalko and Schnitzler (1993) reported a negative correlation between signal duration (or IPI) and signal bandwidth for *Pipistrellus sp.* similar to what we found for *E. fuscus*. The changes in signal duration were mostly due to changes in the tail of the pipistrelle signal. They did not report a lowering of $F_{\min}$ with increasing signal duration. However, the sweep rate at the end of the pipistrelle signal is very low, 0.2 kHz/ms, whereas the sweep rate at the end of *E. fuscus* signals is typically very close to the slope of the regression line in Fig. 3(b), i.e., 0.5 kHz/ms, thus explaining the correlation between $F_{\min}$ and dur in *E. fuscus*. Hence, it is likely, that in the pipistrelle bats, any changes in $F_{\min}$ with signal duration would probably be so small that they would be lost in the overall variation in frequencies emitted by different individuals in different situations.

D. Approach and terminal phase

Our field data showed long duration signals in the approach phase. The duration may exceed 10 ms at the beginning, decreasing gradually to end at 3–4 ms right before the start of the terminal phase. Bats performing in psychophysical experiments usually emit signal with durations of 1–3 ms (Simmons, 1973; Moss and Schnitzler, 1989; Surlykke and Bojesen, 1996). Our data indicate that those signals resemble signals emitted at the transition to or well into the first part of the terminal phase more than they resemble typical approach signals.

The variation in IPI makes it difficult to define the onset of the buzz on the basis of temporal parameters. Long intervals separating the last approach signals from the first buzz signals, as in *P. kuhli* (Schnitzler *et al.*, 1987) did not occur in the *E. fuscus* recordings. However, the change in spectral structure of the signals in these two phases is distinct in all our field recordings. The negative correlation between signal duration and $F_{\min}$ holds all the way through the approach phase, but breaks down when the bats enter the terminal phase. At the transition between the approach and buzz phases of insect pursuit, $F_{\min}$ no longer increases, but instead declines with further decreases in signal duration. Hence, we suggest an exact definition for the onset of the buzz phase: the instance when the correlation between $F_{\min}$ and signal duration becomes positive.

Comparisons of the lab recording with field recordings indicate that neither open-area search phase signals, characterized by relatively long duration ($>$15 ms) shallow FM signals and long IPI, nor search phase signals characteristic of more cluttered areas (call durations around 9–11 ms, IPI's around 120 ms) are produced in the lab. Bats catching insects in the lab operate only in the approach and buzz phases. Our data also indicate that the bats decrease the intensity in the lab, thus corroborating other reports of reduced intensity in confined spaces (Griffin, 1958; Surlykke *et al.*, 1993). Our methods did not allow us to determine the emitted intensities, neither in the field nor in the lab, but an intensity reduction in the lab would explain first that often the signal-to-noise ratio (S/N) of the search/approach signals recorded in the lab was no better than in the field (compare the two 5.4-ms signals in Fig. 5), and second that in spite of the poor S/N the higher frequencies were more prominent in the lab signals.

If we define the start of the terminal buzz using the same criteria as in the field (the first signals where the $F_{\min}$ starts dropping), the duration of the terminal buzz is often surpris-

ingly long in the lab, longer than that recorded in the field. This is somewhat surprising, since the bats have comparatively more knowledge about the position of the prey in the lab (i.e., within the confines of the flight room) than in the field. We do not have a good explanation for this, but it could be that it reflects differences in insect capture strategies in the lab and field. For example, if a bat misses an insect, in the field it may return directly to the search/approach phase in pursuit of a new prey item, but in the lab, the bat may learn that it can turn in flight and catch the same prey item, simply extending the buzz through this adaptive maneuver.

In both the lab and the field, the sound recordings show breaks in signal trains during the approach and buzz phase and a reduction in amplitude during the buzz phase. The high signal-to-noise ratio (S/N) in the lab buzzes helps to confirm the reliability of these two results in the field recordings.

E. Influence of habitat and conspecific bats

The field recordings showed a clear correlation between habitat and search signals, a finding consistent with data reported for other species (Barclay *et al.*, 1999; Kalko and Schnitzler, 1993). We only recorded the long narrow-band search signals in the two open areas (A and B), never in the more cluttered area, C. The mean IPI's decreased according to how open the areas were from A over B and C to the minimum in the lab, probably reflecting how cluttered or confined the bats perceive their surroundings. The very short IPI's and signal durations in the lab indicate that the bats treat the lab flight room as an extremely cluttered habitat.

Our simultaneous recordings of several bats showed greater variation in duration, IPI, and $F_{\min}$ than those of single bats. The directionality of our microphones and of the bat signals makes it likely that bats recorded simultaneously are flying in close proximity. The frequency separation of the signals in the group recordings (for example, those shown in Fig. 4) makes it tempting to speculate that *E. fuscus* exhibits a jamming avoidance response to the signals of other bats in their vicinity (see, e.g., Habersetzer, 1981). However, to directly address the question of jamming avoidance in bats, we need more sophisticated recording methods to track frequency changes in sonar cries with the relative positions and flight directions of two or more individual animals. This could be achieved by mounting lightweight ultrasound microphones with radio transmitters directly on several bats that later forage together, a challenging but not impossible approach. Initial attempts might be taken with established methods, e.g., microphone array recordings (Surlykke *et al.*, 1993) and photographic 3-D reconstruction of flight paths (Kalko and Schnitzler, 1989b).

Fenton (1995) reported geographic variation in *E. fuscus* calls, with bats from eastern North America producing 5–10-ms-long signals and bats from western North America producing 10–15-ms-long signals. However, the variation we have described here from a relatively small geographical area seems to encompass that, and we find it more likely that the reported geographical variation was due to local differences in the habitats where the bats were recorded. Our data indicate that even within the same habitat, bats may change their signals when flying in groups. Our data also indicate that the bats adjust their signals to different flight elevations, in accordance with the results on *E. serotinus* of Jensen and Miller (1999). They found that the higher the bats flew, the longer the duration and the lower the end frequency and bandwidth of the signals. Obrist (1995) reported small, but statistically significant, individual call differences in *E. fuscus*, most often in call frequency. Our data indicate that *E. fuscus* is flexible and adapts its acoustic output (signal duration, repetition rate, and spectral content) to habitat, flight elevation, and/or the presence of other bats. If these parameters are taken into account, it is likely that individual call design differences would become more apparent (see Masters *et al.*, 1995).

V. CONCLUSIONS

Our recordings illustrate echolocation signals produced by *E. fuscus* hunting insects under natural and laboratory conditions. The field recordings show distinct patterns of signals recorded under different hunting conditions and during different phases of insect pursuit. The data also point to potentially important variations in sonar signals not observed in the laboratory, and it is likely that more recordings would reveal even greater flexibility in sonar production patterns in this species. We believe that the acoustic data presented here provide a useful frame of reference for both behavioral and neurophysiological laboratory studies of echolocation in *E. fuscus*.

ACKNOWLEDGMENTS

This work was supported by The Danish Natural Research Foundation and Danish Natural Science Research Council (A.S.), the National Science Foundation (Grant No. IBN-9258255), the Whitehall Foundation (Grant No. S97-20), and the National Institutes of Health (Grant No. R01-MH56366) (C.F.M.). We thank Shiva Sinha and Pete Abrams for assistance in the field, and Amy Akryjak, Hannah Gilkenson, and Aaron Schurger for assistance in the lab.

Barclay, R. M. R., Fullard, J. H., and Jacobs, D. S. (**1999**). ''Variation in the echolocation calls of the hoary bat (*Lasiurus cinereus*): Influence of body size, habitat, structure, and geographic location,'' Can. J. Zool. **77**, 530–534.

Betts, B. J. (**1998**). ''Effects of interindividual variation in echolocation calls on identification of big brown and silver-haired bats,'' J. Wildlife Manage. **62**, 1003–1010.

Covey, E., and Casseday, J. H. **1995**. ''The lower brainstem auditory pathways,'' in *Hearing by Bats*, edited by A. N. Popper and R. R. Fay (Springer, New York), Chap. 6, pp. 235–295.

Fenton, M. B. (**1986**). ''Design of bat echolocation calls: Implications for foraging ecology and communication,'' Mammalia **50**, 193–204.

Fenton, M. B. (**1995**). ''Natural history and biosonar signals,'' in *Hearing by Bats*, edited by A. N. Popper and R. R. Fay (Springer, New York), Chap. 2, pp. 37–86.

Griffin, D. R. (**1944**). ''Echolocation in blind men, bats, and radar,'' Science **100**, 589–590.

Griffin, D. R. (**1958**). *Listening in the Dark*, 2nd ed. (Yale University Press, New Haven).

Habersetzer, J. (**1981**). ''Adaptive echolocation sounds in the bat *Rhinopoma hardwickei*,'' J. Comp. Physiol. A **144**, 559–566.

Hartley, D. J. (**1992**). ''Stabilization of perceived echo amplitudes in echolocating bats. II. The acoustic behavior of the big brown bat, *Eptesicus fuscus*, when tracking moving prey,'' J. Acoust. Soc. Am. **91**, 1133–1149.

Hartley, D. J., and Suthers, R. A. (**1989**). ''The sound emission pattern of the echolocating bat, *Eptesicus fuscus*,'' J. Acoust. Soc. Am. **85**, 1348–1351.

Jensen, M. E., and Miller, L. A. (**1999**). ''Echolocation signals of the bat *Eptesicus serotinus* recorded using a vertical microphone array: Effect of flight altitude on searching signals,'' Behav. Ecol. Sociobiol. **47**, 60–69.

Kalko, E. K. V. (**1994**). ''Coupling of sound emission and wingbeat in naturally foraging European pipistrelle bats (Microchiroptera: Vespertilionidae),'' Folia Zool. **43**, 363–376.

Kalko, E. K. V., and Schnitzler, H.-U. (**1989a**). ''Two-wave-front interference patterns in frequency-modulated echolocation signals of bats flying low over water,'' J. Acoust. Soc. Am. **85**, 961–962.

Kalko, E. K. V., and Schnitzler, H.-U. (**1989b**). ''The echolocation and hunting behavior of Daubenton's bat, *Myotis daubentoni*,'' Behav. Ecol. Sociobiol. **24**, 225–238.

Kalko, E. K. V., and Schnitzler, H.-U. (**1993**). ''Plasticity in echolocation signals of European pipistrelle bats in search flight: Implications for habitat use and prey detection,'' Behav. Ecol. Sociobiol. **33**, 415–428.

Lawrence, B. D., and Simmons, J. A. (**1982**). ''Measurements of atmospheric attenuation at ultrasonic frequencies and the significance for echolocation by bats,'' J. Acoust. Soc. Am. **71**, 585–590.

Masters, W. M., Raver, K. A. S., and Kazial, K. A. (**1995**). ''Sonar signals of big brown bats, *Eptisicus fuscus*, contain information about individual identity, age and family affiliation,'' Anim. Behav. **50**, 1243–1260.

Moss, C. F., and Schnitzler, H.-U. (**1989**). ''Accuracy of target ranging in echolocating bats: Acoustic information processing,'' J. Comp. Physiol. A **165**, 383–393.

Moss, C. F., and Schnitzler, H.-U. (**1995**). ''Behavioral studies of auditory information processing,'' in *Hearing by Bats*, edited by A. N. Popper and R. R. Fay (Springer, New York), Chap. 3, pp. 87–145.

Neuweiler, G. (**1984**). ''Foraging, echolocation and audition in bats,'' Naturwissenschaften **71**, 446–455.

Obrist, M. K. (**1995**). ''Flexible bat echolocation: The influence of individual, habitat and conspecifics on sonar signal design,'' Behav. Ecol. Sociobiol. **36**, 207–219.

Popper, A. N., and Fay R. R. (**1995**). *Hearing by Bats* (Springer, New York).

Schnitzler, H.-U., and Kalko, E. K. V. (**1998**). ''How echolocating bats search and find food,'' in *Bat Biology and Conservation*, edited by T. H. Kunz and P. A. Racey (Smithsonian Institution, Washington), Chap. 12, pp. 183–196.

Schnitzler, H.-U., Kalko, E. K. V., Miller, L. A., and Surlykke, A. (**1987**). ''The echolocation and hunting behavior of the bat, *Pipistrellus kuhli*,'' J. Comp. Physiol. A **161**, 267–274.

Simmons, J. A. (**1973**). ''The resolution of target range by echolocating bats,'' J. Acoust. Soc. Am. **54**, 157–173.

Simmons, J. A. (**1987**). ''Acoustic images of target range in the sonar of bats,'' Nav. Res. Rev. **39**, 11–26.

Simmons, J. A., Fenton, M. B., and O'Farrell, M. J. (**1979**). ''Echolocation and pursuit of prey by bats,'' Science **203**, 16–21.

Simmons, J. A., Ferragamo, M., Saillant, P., Haresign, T., Wotton, J. M., Dear, S. P., and Lee, D. N. (**1995**). ''Auditory dimensions of acoustic images in echolocation,'' in *Hearing by Bats*, edited by A. N. Popper and R. R. Fay (Springer, New York), Chap. 4, pp. 146–190.

Simmons, J. A., Howell, D. J., and Suga, N. (**1975**). ''Information content of bat sonar echoes,'' Am. Sci. **63**, 204–215.

Simmons, J. A., and Stein, R. A. (**1980**). ''Acoustic imaging in bat sonar: Echolocation signals and the evolution of echolocation,'' J. Comp. Physiol. A **135**, 61–84.

Speakman, J. R., and Racey, P. (**1991**). ''No cost of echolocation for bats in flight,'' Nature (London) **350**, 421–423.

Suga, N. (**1988**). ''What does single-unit analysis in the auditory cortex tell us about information processing in the auditory system?,'' in *Neurobiology of Neocortex*, edited by P. Rakic and W. Singer (Wiley, New York), pp. 331–350.

Surlykke, A., and Bojesen, O. (**1996**). ''Integration time for short broad band clicks in echolocating FM-bats (*Eptesicus fuscus*),'' J. Comp. Physiol. A **178**, 235–241.

Surlykke, A., Miller, L. A., Møhl, B., Andersen, B. B., Christensen-Dalsgaard, J., and Jørgensen, M. B. (**1993**). ''Echolocation in two very small bats from Thailand: *Craseonycteris thonglongyai* and *Myotis siligorensis*,'' Behav. Ecol. Sociobiol. **33**, 1–12.

Wilson, W. W., and Moss, C. F. (**2000**). ''Sensory-motor behavior of free-flying FM bats during target capture,'' in *Advances in the Study of Echolocation in Bats and Dolphins*, edited by J. A. Thomas, C. F. Moss, and M. Vater (Chicago University Press, Chicago, 2000), Chap. 1.3.

The single sonic muscle twitch model for the sound-production mechanism in the weakfish, *Cynoscion regalis*

Mark W. Sprague[a)]
Department of Physics, East Carolina University, Greenville, North Carolina 27858

(Received 9 May 2000; accepted for publication 14 August 2000)

A model of the weakfish, *Cynoscion regalis*, sound-production mechanism based on damped driven oscillators is presented. The weakfish ''purr'' consists of several short pulses of sound separated by intervals of no sound. Each pulse is produced by an individual simultaneous twitch of each sonic muscle causing swimbladder oscillations that radiate sound into the surrounding water. The sonic muscles are modeled as a stretched string with a time-varying tension force. The swimbladder is modeled as a highly damped, driven oscillator undergoing radial oscillations. Although the swimbladder functions as an impedance-matching device between the sonic muscles and the surrounding fluid, its transient response to the sonic muscle pulses includes frequencies that are not a part of the spectrum of the sonic muscle excitations. Differential equations of motion for the sonic muscles and swimbladder are given and are numerically solved to produce predicted waveforms matching those of measured weakfish sounds. The model leads to better understanding of the weakfish sound-producing mechanism and the effects of environmental and physiological factors on sound production. This model may also lead to a better understanding of sound production in other species, particularly other members of Family Sciaenidae, with similar mechanisms for sound production. [S0001-4966(00)03711-5]

PACS numbers: 43.80.Ka, 43.30.Sf, 43.40.At [WA]

I. INTRODUCTION

A. Sound production in fishes

Many fishes make species-specific sounds to communicate with one another (Myrberg *et al.*, 1965; Fish and Mowbray, 1970; Fine *et al.*, 1977; Myrberg, 1981; Mann *et al.*, 1997). These sounds are associated with aggression, aggregation, fright, and reproductive behaviors (Winn, 1964). Males of the Family Sciaenidae make species-specific drumming sounds during courtship of the females at locations where spawning occurs (Fish and Mowbray, 1970; Mok and Gilmore, 1983; Connaughton and Taylor, 1995; Connaughton and Taylor, 1996; Luczkovich *et al.*, 1999a). The spawning-related sounds produced by Sciaenids have been used to identify and map areas where spawning occurs (Luczkovich *et al.*, 1999a, 1999b, 2000).

B. Weakfish sound-production mechanism

Many soniferous fishes produce sound by driving their swimbladders with sonic muscles (Smith, 1905; Tower, 1908; Demski *et al.*, 1973). Some sounds produced by sonic muscle/swimbladder mechanisms are continuous tones resulting from repeated sonic muscle twitches driving the swimbladder in a steady-state response. Examples of these continuous-tone sounds are the oyster toadfish *Opsanus tau* ''boatwhistle'' (Skoglund, 1961; Winn, 1964; Fish and Mowbray, 1970; Rome *et al.*, 1996), the northern midshipman *Porichthys notatus* ''buzz'' and ''hum'' (Cohen and Winn, 1967; Bass and Marchaterre, 1989), and the myomyrid electric fish *Pollymyrus isidori* ''moan'' (Crawford *et al.*, 1986). Another class of sonic muscle/swimbladder sounds is a pulse in which one or more sonic muscle twitches drives the swimbladder in a transient response. Many species produce trains of repeated pulses of sound with each pulse decaying before the next pulse begins. Examples of pulse-trains are the oyster toadfish ''grunt'' (Winn, 1964; Fish and Mowbray, 1970), the northern midshipmen ''grunt'' (Cohen and Winn, 1967; Bass and Marchaterre, 1989), the northern searobin *Prionotus carolinus* ''squawk,'' ''bark,'' and ''growl'' (Fish and Mowbray, 1970), and the drumming sounds produced by members of Family Sciaenidae such as weakfish *Cynoscion regalis* (Smith, 1905; Tower, 1908; Fish and Mowbray, 1970; Sprague *et al.*, 2000), spotted seatrout *C. nebulosus* (Mok and Gillmore, 1983; Sprague *et al.*, 2000), red drum *Sciaenops ocellatus* (Fish and Mowbray, 1970; Guest and Lasswell, 1978; Sprague *et al.*, 2000), silver perch *Bairdiella chrysoura* (Mok and Gillmore, 1983; Sprague *et al.*, 2000), and Atlantic croaker *Micropogonias undulatus* (Fish and Mowbray, 1970; Demski *et al.*, 1973).

The swimbladder is an air-filled sac in the body cavity of most fishes with primary function to control buoyancy (Fig. 1). The swimbladder as a sound-producing mechanism has been studied in many other species in addition to weakfish. Tavolga (1964) discovered that when he released gas from an oyster toadfish swimbladder, the frequency content of the oscillation did not change but the amplitude of the sound decreased. He concluded that the swimbladder functioned as an impedance-matching device driven by the sonic muscles and that the resonant properties of the swimbladder were not as important to the sound produced as the driving force produced by the sonic muscles. This is generally true for continuous-tone sounds, but the acoustic properties of both the sonic muscles and swimbladder are important in

a)Electronic mail: spraguem@mail.ecu.edu

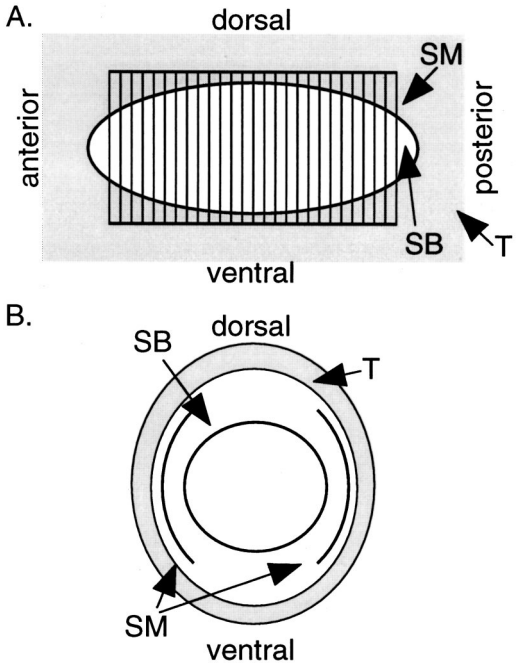


FIG. 1. A simplified view of a weakfish swimbladder and sonic muscles. For a more detailed picture, see Ono and Poss (1982) or Connaughton and Taylor (1995). SM, sonic muscles; SB, swimbladder; T, surrounding tissue. (a) Side view. (b) Cross-sectional view.

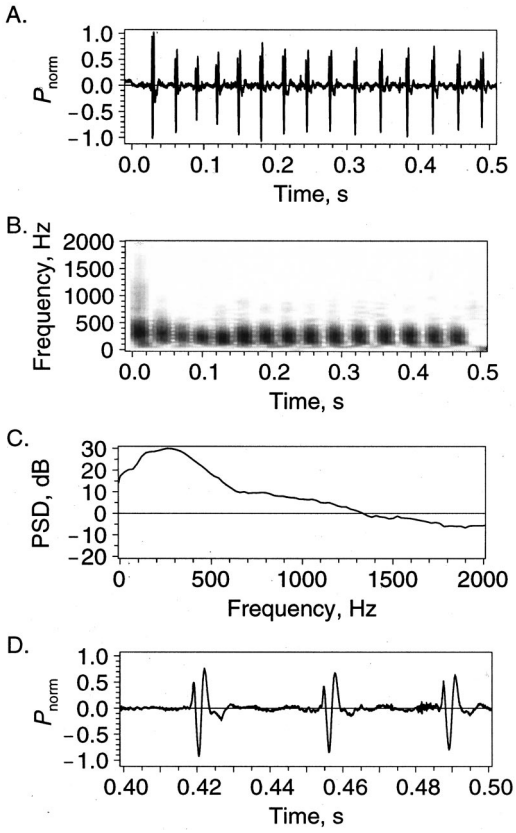


FIG. 2. A ''purr'' produced by a captive weakfish (340-mm standard length) in an 89-L cooler filled with seawater. The sound was captured using an InterOcean Model T-902 hydrophone (frequency range 20–10 000 Hz, sensitivity -195 dB *re* 1 V/μPa) connected to an amplifier (gain adjustable from 15 to 95 dB). The amplifier output was recorded by a Sony TCD-D8 digital audio tape (DAT) recorder (frequency range 20–22 000 Hz$\pm$1 dB) with 16 bits of resolution and a sampling rate of 48 kHz. The sound was resampled at 24 kHz using a National Instruments NB-2150F analog-to-digital board with anti-aliasing filters. Low-frequency noise below 70.3 Hz was removed using a Fourier filter. (a) An oscillogram of the entire purr. The vertical axis is normalized pressure (P_{norm}). (b) A sonogram of the entire purr computed using fast Fourier transforms (FFTs) of 1024-point Hanning windows spaced every 128 points in the sampled waveform. The gray scale represents the power spectral density (PSD) on a logarithmic scale with darker tones representing higher PSDs. (c) An average power spectrum of the entire purr. The power spectra in the sonogram from (b) were averaged, and the PSDs were converted to decibels. (d) A magnified oscillogram of the purr from (a).

determining the frequencies in pulse-train sounds because the swimbladder continues to oscillate after the sonic muscle twitches stop.

Harris (1964) showed that a swimbladder radiating sound could be approximated as an oscillating sphere, because the swimbladder radius was much smaller than the wavelength of the radiated sound. He concluded that monopole radiation caused by volume oscillations of swimbladders was the most efficient process of sound radiation.

Fishes produce swimbladder vibrations using sonic muscles. The sonic muscles of oyster toadfish are intrinsic to the swimbladder—or connected to the swimbladder walls—and occur on both sides of the swimbladder (Fine, 1997). Oyster toadfish sonic muscles are capable of repeated contractions of less than 10 ms each (Tavolga, 1964; Rome *et al.*, 1996), and the rate of the muscle contractions determines the fundamental frequency of the radiated sound (Skoglund, 1961). Weakfish sonic muscles evolved from different structures than those of the oyster toadfish (Fine, 1997). Male weakfish have sonic muscles on both sides of their swimbladders (Fig. 1), but females have no sonic muscles. Sonic muscles in weakfish are extrinsic to the swimbladder and are not attached to the swimbladder walls (Tower, 1908; Fine, 1997; Connaughton *et al.*, 2000). The muscles extend the length of the swimbladder with fibers running dorsoventrally from an attachment at aponeurotic sheet on the dorsal side of the swimbladder to an attachment at the hypaxial muscles on the ventral side (Ono and Poss, 1982; Hill *et al.*, 1987).

The weakfish sound, known as a ''purr,'' consists of several pulses of high-frequency oscillation within a lower-frequency envelope. The average dominant frequency in field recordings of weakfish purrs in a study in North Carolina waters was 347 Hz (Sprague *et al.*, 2000). Figure 2 shows an oscillogram, a sonogram, an average power spectrum, and a magnified oscillogram of a purr produced by a captive male weakfish of 340-mm standard length (Sprague *et al.*, 2000). In the first oscillogram [Fig. 2(a)] of the entire purr there were 15 pulses within a 0.5-s interval. The sonogram [Fig. 2(b)] and average power spectrum [Fig. 2(c)] show that the purr had a broad frequency peak near 300 Hz. The magnified oscillogram [Fig. 2(d)] shows that each pulse had three high-amplitude half-cycles sometimes followed by decaying lower amplitude oscillations. The individual pulses in weakfish purrs are about 3.5 ms in duration with interpulse intervals of 30–50 ms (Connaughton *et al.*, 2000; Sprague *et al.*, 2000). Connaughton *et al.* (2000) measured simultaneous weakfish acoustic signals and sonic muscle electromyograms (EMGs) and determined that each pulse in the acoustic signal was preceded by simultaneous single twitches

of the sonic muscles on both sides of the fish. The EMGs of weakfish sonic muscle twitches ranged from 7.9–13.6 ms in duration. Each muscle twitch produced multiple vibrations of the swimbladder causing the radiated sound to have different frequency content than the muscle twitch. These findings indicated that the weakfish sound-producing system was not fully understood. The weakfish sonic muscle and swimbladder system is apparently different from the toadfish sound-producing system in which the frequency of the muscle twitch determines the frequency of the swimbladder oscillations (Skoglund, 1961).

C. SSMT model proposed

A physical model of the weakfish sound-production mechanism will allow scientists to understand the physical parameters required to produce the measured weakfish sounds and give research biologists insight into biophysical studies involving the sound-production mechanism in weakfish and other species. Such a model will also give insight into the sonic muscle and swimbladder physiology that produces this sound. The single sonic muscle twitch (SSMT) model is based on measurements of the weakfish sonic muscle EMGs (Connaughton *et al.*, 2000) and its relationship to the sound waves radiated by the swimbladder.

The SSMT model assumes the weakfish sonic muscles vibrate as a stretched string with time-varying tension force. When the sonic muscle is in its relaxed state, the tension is very small, and the muscle does not vibrate. As the muscle tension rapidly increases, it is suddenly out of equilibrium and begins to oscillate—much like a plucked string on a guitar—until it relaxes again, causing the oscillations to stop. The resonant frequency of the sonic muscle changes during a twitch in much the same way as the resonant frequency of a guitar string changes as its tension increases and decreases. Oscillations of the sonic muscle and swimbladder are weakly coupled because the weakfish sonic muscle is not attached to the swimbladder. This weak coupling allows the sonic muscle oscillations to be affected very little by the swimbladder much the same as an amplifier signal output is affected very little by a connected loudspeaker.

The swimbladder in the SSMT model is a highly damped oscillator driven by the sonic muscle vibrations. The swimbladder vibrates as a damped, driven oscillator and acts as an impedance-matching device between the sonic muscle and the surrounding water just as a guitar body radiates the vibrations of the strings or a loudspeaker radiates an amplifier signal.

II. THEORY

A. Damped oscillator

A mechanical oscillating system with velocity-dependent linear damping force $-b\dot{x}$ is described by the well-known differential equation (Marion and Thornton, 1995),

$$\ddot{x}(t)+2\beta\dot{x}(t)+\omega_0^2 x(t)=\frac{F(t)}{m}, \tag{1}$$

where $x(t)$ is the displacement, $\beta=b/(2m)$ the damping parameter, $\omega_0=2\pi f_0$ the angular frequency for undamped oscillations, $F(t)$ the driving force, and m mass. The dots over x represent derivatives with respect to time. The parameter b describes the ratio of the damping force exerted on the mass by the surrounding fluid to the velocity of the mass. This damping ratio depends on the shape of the mass and the ability of the fluid to dissipate the energy and momentum of the mass. The damping ratio is usually determined experimentally by measuring the amplitude decay of the oscillations of the mass and comparing it to the solution of Eq. (1). When $\beta<\omega_0$ the system is underdamped and transient motion consists of sinusoidal oscillation in an exponentially decaying envelope. If the driving force is periodic, the steady-state behavior of the oscillator will have the same frequency content as the driving force. When $\beta>\omega_0$, the system is overdamped. The transient displacement is not oscillatory but exponentially decays approaching equilibrium.

B. Stretched string

A stretched string with length L, tension force T, and mass per unit length μ oscillates with fundamental frequency f_0 given by (Marion and Thornton, 1995)

$$f_0=\frac{1}{2L}\sqrt{\frac{T}{\mu}}. \tag{2}$$

Thus the resonant frequency of the string is proportional to the square root of the string tension force and inversely proportional to the square root of the mass per unit length of the string. As the tension force increases, the resonant frequency increases.

The displacement of a string vibrating at its fundamental frequency depends on the position y along the string and time t. If the string has length L and velocity-dependent damping ratio b_s, the displacement x is (Marion and Thornton, 1995)

$$x(y,t)=\eta(t)\sin\left(\frac{\pi y}{L}\right). \tag{3}$$

The sinusoidal displacement along the y direction oscillates according to time dependence $\eta(t)$, which is a solution of the differential equation

$$\ddot{\eta}(t)+\beta_s\dot{\eta}(t)+\omega_0\eta(t)=0. \tag{4}$$

Equation (4) is similar to Eq. (1) for a damped oscillator except that in Eq. (4) $\beta_s=b_s/(2\mu)$ and $\omega_0=2\pi f_0$ is determined by Eq. (2).

C. Sound radiated by a vibrating sphere

A sphere undergoing volume oscillations radiates sound as a monopole with amplitude inversely proportional to the distance r from the sphere center (Pierce, 1989). The far-field sound pressure p due to a radially oscillating sphere of radius a for low frequencies (determined by the condition that the sound wave travels many sphere radii during a single cycle, $f\ll c/a$) is

$$p(r,t)=\frac{\rho}{4\pi r}\left(\frac{dQ_s}{dt}\right)_{t\to t-r/c}. \tag{5}$$

In Eq. (5), r is the radial distance from the center of the sphere; t is time; ρ is the fluid density; c is the sound speed, and $Q_s=4\pi a^2\nu_s$ is the volume velocity of the sphere, where ν_s is the radial velocity of the sphere wall. The parameter $t-r/c$ refers to the time at which the wave would have originated at the sphere center. Thus the radiated sound pressure is proportional to the time derivative of the volume velocity. Although Eq. (5) assumes that the sphere is in a uniform unbounded medium, it can readily be adapted to other environments by accounting for the diffraction and reflection of the waves due to changes or boundaries in the medium.

III. MODEL

A. Sonic muscle

The sonic muscle begins in a relaxed state with no tension but in a position that would be out of equilibrium if there were a tension force. As the sonic muscle undergoes a contraction and subsequent relaxation, its resonant frequency increases and decreases in proportion to the square root of the muscle tension as given by Eq. (2). The use of the term ''resonant frequency'' of the sonic muscle is misleading, as the sonic muscle does not oscillate at resonance during its single twitch. The resonant frequency represents that frequency at which the sonic muscle would oscillate if the tension force were held constant with no damping. I used the sonic muscle resonant frequency as an adjustable parameter in the model rather than the tension force because the resonant frequency appears directly in Eq. (4). When the sonic muscle contracts, the tension force pulls it toward the equilibrium position, and it begins underdamped oscillation. The increasing muscle tension adds potential energy to the oscillating system when the sonic muscle is displaced from equilibrium. As the muscle relaxes, the decreasing tension removes potential energy from the system when the sonic muscle is displaced from equilibrium. When the resonant angular frequency of the muscle becomes less than the damping parameter β_m, the system becomes overdamped, and the oscillations cease.

I modeled the increasing and decreasing sonic muscle resonant frequency as two Gaussian functions given by

$$f_{0m}(t)=\begin{cases} f_{\max}e^{-\gamma_1(t-t_0)^2}, & t<t_0 \\ f_{\max}e^{-\gamma_2(t-t_0)^2}, & t>t_0, \end{cases} \tag{6}$$

where t_0 is the time of the peak muscle tension. The parameters γ_1 and γ_2 determine the time for a sonic muscle twitch. The maximum sonic muscle resonant frequency, $f_{\max}$, depends on the maximum sonic muscle tension as well as the sonic muscle length and mass according to Eq. (2). I used the values of γ_1, γ_2, and $f_{\max}$ as adjustable parameters. Figure 3 shows a plot of the sonic muscle resonant frequency during a twitch.

The sonic muscle oscillates as a damped stretched string according to Eqs. (3) and (4). The equation of motion for the time dependence of the sonic muscle displacement is

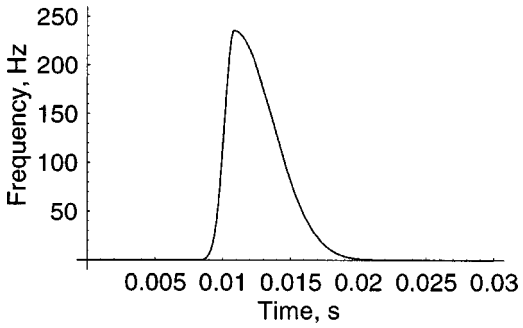


FIG. 3. The resonant frequency of the modeled weakfish swimbladder versus time. The resonant frequency is given by Eq. (6) using the values for $f_{\max}$, γ_1, and γ_2 given in Table I.

$$\ddot{\eta}(t)+2\beta_m\dot{\eta}(t)+[\omega_{0m}(t)]^2\eta(t)=0, \tag{7}$$

where $\omega_{0m}(t)=2\pi f_{0m}(t)$ is given by Eq. (6). The initial displacement, $\eta(0)$, is nonzero, and the initial velocity, $\dot{\eta}(0)$, is zero. The damping parameter for the sonic muscle, which determines the decay rate of the sonic muscle oscillations, is $\beta_m=b_m/(2\mu_m)$, where b_m is the velocity-dependent damping ratio and μ_m the sonic muscle mass per unit length. This system is nonlinear due to the time dependence of ω_{0m}. Equation (7) has no closed-form, exact solution, but numerical techniques may be used to find a high-precision approximate solution.

B. Swimbladder

The swimbladder is an oscillator with a high damping parameter in which transient behavior decays rapidly causing it to follow the driving force of the sonic muscle. The equation of motion for the swimbladder is that of a damped, driven oscillator,

$$\ddot{x}_b(t)+2\beta_b\dot{x}_b(t)+\omega_{0b}x_b(t)=F_{\text{ext}}(t)/m_b, \tag{8}$$

where $x_b(t)$ is the volume displacement, $\beta_b=b/(2m_b)$ the damping parameter, $\omega_{0b}=\sqrt{1/(m_bC_A)}=2\pi f_{0b}$ the resonant angular frequency for undamped oscillations, m_b the effective swimbladder mass, C_A the swimbladder acoustic compliance, and f_{0b} the swimbladder resonant frequency. The driving force $F_{\text{ext}}(t)$ is due to the sonic muscle and is proportional to the sonic muscle acceleration, $\ddot{\eta}(t)$ given by Eq. (7). Although Eq. (8) is linear, it must be solved numerically, as the driving force is an approximate function obtained from the numerical solution of Eq. (7). Due to Eq. (5), the sound radiated by the swimbladder is proportional to the swimbladder volume acceleration, $\ddot{x}_b(t)$.

C. Equivalent circuit

Figure 4 shows a diagram of an equivalent electrical circuit for the SSMT-model weakfish sound-producing mechanism. Each element in the circuit represents an equivalent term in Eqs. (7) or (8). The parameter R_{trans} represents transmission loss between the sonic muscles and the swimbladder. As the radiated sound is proportional to the swimbladder volume acceleration, it is analogous to the voltage across the swimbladder inductor, which supplies the input voltage for the water element in the circuit.

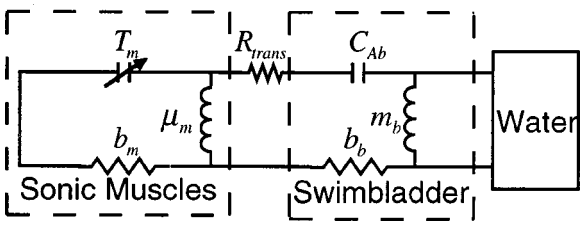


FIG. 4. Equivalent electrical circuit to the weakfish sound-production mechanism. The values for each element in the circuit represent the equivalent terms in Eqs. (7) and (8). The parameter R_{trans} is due to the transmission loss between the sonic muscle and the swimbladder. The acoustic pressure of the radiated sound is analogous to the voltage across the water element in the circuit.

IV. RESULTS

I solved Eqs. (7) and (8) numerically using the NDSolve function in the MATHEMATICA program (Wolfram, 1996). I adjusted the parameters $f_{\max}$, γ_1, γ_2, β_m, f_{0b}, and β_b so that the predicted waveform matched the measured waveform. The parameters for which the predicted waveform best fit the measured waveform are given in Table I. Figure 5 shows the motion of the modeled sonic muscles during a single twitch. At the end of the twitch, the sonic muscles return to the same position that they were in at the beginning of the twitch, preparing them for the start of the next twitch. Figure 6 shows a comparison of normalized measured and predicted waveforms for a single pulse. The waveform predicted by the SSMT model is in agreement with the measured waveform. Figure 7 shows the real and imaginary parts of a fast Fourier transform (FFT) of the measured and predicted waveforms, and Fig. 8 shows power spectra of the same two signals. In each case the measured and predicted curves are in agreement.

V. DISCUSSION

The SSMT model demonstrates how weakfish produce their sound pulses with single sonic muscle twitches. Although weakfish and oyster toadfish both use their swimbladders as impedance-matching devices, weakfish sound production is quite different from that of the oyster toadfish. Weakfish, unlike toadfish, produce short pulses of sound separated by intervals of no sound. The extrinsic sonic muscles of the weakfish are not attached to the swimbladder, allowing the sonic muscles to vibrate separately from the swimbladder. Each muscle twitch drives the swimbladder, which radiates sound into the surrounding water. At the end of each twitch, the sonic muscles return to their original position to be ready for the next twitch.

TABLE I. Values used for best fit between the predicted and measured waveforms.

Parameter	Value
$f_{\max}$	235.7 Hz
γ_1	1.111×10^{6} s^{-2}
γ_2	6.050×10^{-4} s^{-2}
β_m	324.8 s^{-1}
β_b	1248 s^{-1}
f_{0b}	209.0 Hz

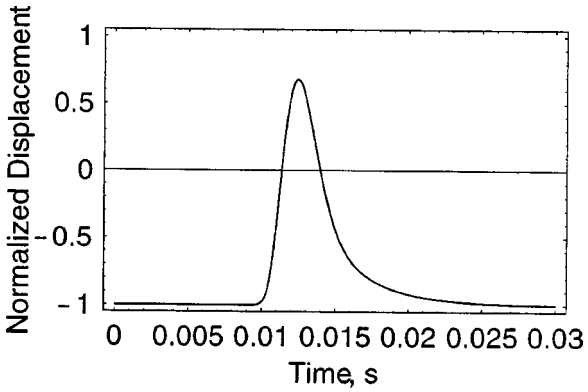


FIG. 5. The motion of the sonic muscles during a single twitch. This curve was obtained by numerical solution of Eq. (7) using the parameters given in Table I.

A. Model assessment

The agreement between the waveform predicted by the SSMT model and the measured waveform suggests that the weakfish sound-producing mechanism consists of a vibrating, stretched sonic muscle that causes swimbladder oscillations. The model produces the correct waveform from a single muscle twitch, in agreement with EMG measurements (Connaughton *et al.*, 2000). While weakfish sonic muscle contractions have not been measured, the muscle twitch times suggested by the SSMT model (see Fig. 3) indicate that weakfish sonic muscle contractions may be as fast as those of the toadfish (Rome *et al.*, 1996). Slight differences between the predicted and measured waveforms may be caused by the muscle twitch being different than the Gaussian time dependence assumed here. Alternatively, the swimbladder damping parameter may have increased suddenly at the end of the pulse. A possible cause for an increase in the swimbladder damping parameter is contact with the relaxed sonic muscles. The differences between the measured and predicted waveforms only lead to minor differences in the FFTs and power spectra between the measured and predicted sounds as shown in Figs. 7 and 8.

Connaughton *et al.* (2000) measured variations in the amplitude of successive sound pulses while the EMGs of

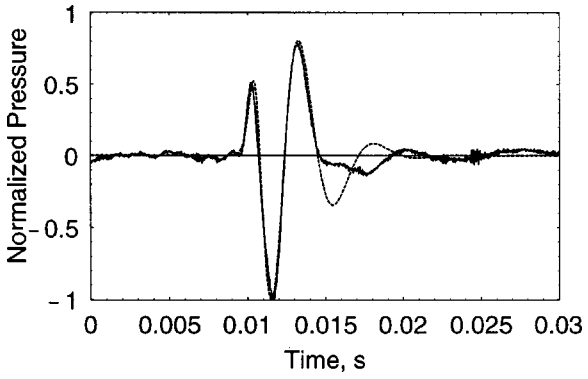


FIG. 6. Predicted and measured waveforms for a weakfish purr. The solid curve is the measured purr, which was Fourier filtered to remove low-frequency noise below 70.3 Hz. The dashed curve is the purr given by the numerical solution of Eq. (8). Both waveforms were normalized so the largest displacement was ±1. The parameters used to calculate the numerical solution are given in Table I.

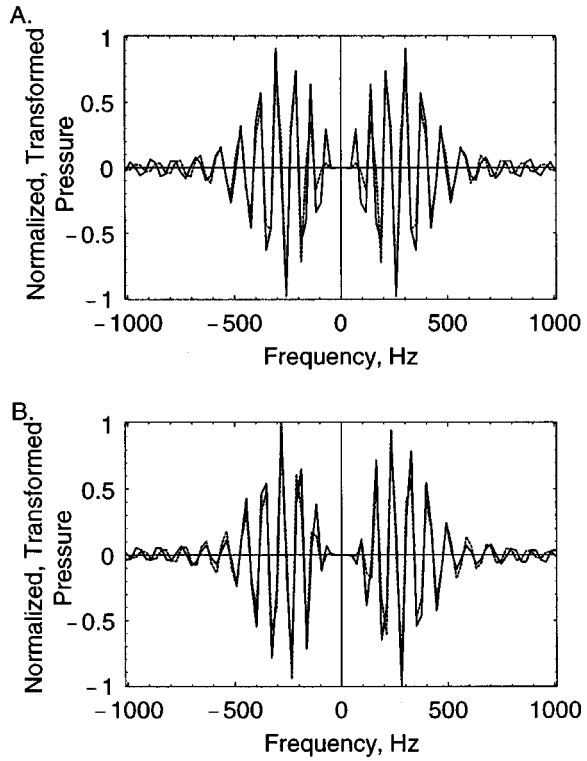


FIG. 7. Predicted and measured fast Fourier transforms (FFTs) of the weakfish purr. The solid curve is an FFT of the measured purr, and the dashed curve is an FFT of the waveform predicted by the SSMT model. The parameters used to calculate the predicted purr are given in Table I. The FFTs were calculated using 1024 points with a sampling frequency of 24 000 Hz. No window function was used, as each 1024-point FFT contained the entire pulse waveform. Both FFT's were normalized so that the largest frequency component was equal to ± 1. (a) The real part of the FFT. (b) The imaginary part of the FFT.

each muscle twitch were almost identical. The SSMT model can explain this observation if the sonic muscle begins successive pulses at different initial positions causing different amplitude oscillations produced from the identical muscle twitches. Another factor that could cause different amplitude

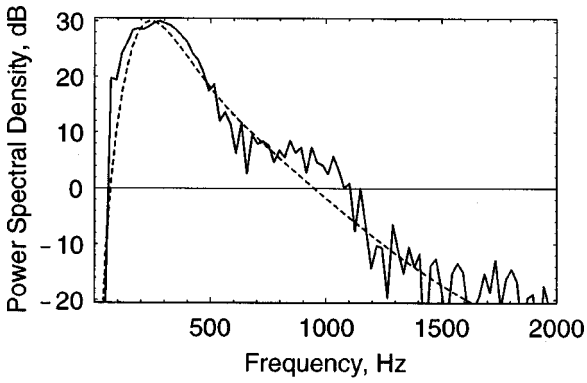


FIG. 8. Predicted and measured power spectra of the weakfish purr. Each power spectrum was calculated from the FFT shown if Fig. 7. The solid curve is a power spectrum of the measured purr, and the dashed curve is a power spectrum of the waveform predicted by the SSMT model. The parameters used to calculate the predicted purr are given in Table I. The FFTs were calculated using 1024 points with a sampling frequency of 24 000 Hz. No window function was used, as each 1024-point FFT contained the entire pulse waveform. Both power spectra were normalized so the peak power spectral density was 30 dB.

sound pulses with the same EMG is that different numbers of motor groups could be recruited by the same nerve during successive muscle contractions.

The experimental weakfish purr used to fit the SSMT model was measured in a small (89-L) rectangular box-shaped container rather than in a free-field environment. The largest dimension of the container was 0.775 m, corresponding to a minimum normal mode frequency of 980 Hz. Since all of the significant frequency content of the measured purr (see Fig. 8) was much lower than 980 Hz, the container had little effect on the sound. Nevertheless, the model parameters could easily be adjusted to predict a slightly different weakfish waveform measured in a free-field environment.

The fit parameters given in Table I are related to measurable physiological and acoustical characteristics. The maximum sonic muscle resonant frequency, $f_{\max}$, depends on the maximum muscle tension as well as the sonic muscle mass and length according to Eq. (2). While the mass and size of the weakfish sonic muscle have been measured (Ono and Poss, 1982; Hill *et al.*, 1987), the sonic muscle tension has not. The parameters γ_1 and γ_2 determine the time dependence of the sonic muscle twitch and could be fit to a measurement of the sonic muscle tension variation in time. The sonic muscle damping parameter, β_m, could be fit to a measurement of the sonic muscle tension and displacement. The swimbladder resonant frequency, f_{0b}, and damping parameter, β_b, could be measured experimentally using a sound source to determine the target strength versus frequency curve of the swimbladder (Løvik and Hovem, 1979). The resonant frequency is the maximum of the target strength versus frequency curve, and the damping parameter determines the width of the peak (Marion and Thornton, 1995).

B. Variations in weakfish sound production

The SSMT model may lead to a better understanding of the effects of fish size and environmental factors like temperature and pressure on weakfish sounds. The model suggests that the waveform depends on the properties of the sonic muscle as well as the resonant frequency of the swimbladder. The sonic muscle driving force determines the shape of the first two half-cycles in the weakfish purr waveform predicted by the SSMT model, while the swimbladder resonant frequency and damping parameter determine the shape of the subsequent oscillations. Variations in the frequency content of sounds produced by different individual weakfish may be caused by differences in the sonic muscle mass and strength between them in addition to differences in swimbladder size.

Decreasing the sonic muscle resonant frequency decreases the frequency and increases the period of the first two half-cycles in the waveform predicted by the model. If larger individuals have longer and more massive sonic muscles they would have lower sonic muscle resonant frequencies at the same tension because of Eq. (2). This difference could be offset if the larger sonic muscles were capable of higher tensions, increasing their resonant frequencies. The prediction of decreasing frequency with increasing fish size is consistent with the frequency versus size measurements of Con-

naughton *et al.* (2000). Seasonal variations in weakfish sounds may be caused by changes in the sonic muscle properties of individual fish as the spawning season progresses. Connaughton *et al.* (1997) measured changes in weakfish sonic muscle fiber cross-sectional area and muscle protein content along with a decrease in the dominant frequency of the sound produced as the sonic muscles changed from a hypertrophied state during the late spring spawning period to an atrophied state in the late summer. The SSMT model would predict the same seasonal differences in the waveforms of weakfish sounds due to these muscle changes, as sonic muscle hypertrophy would lead to higher sonic muscle tensions and resonant frequencies and sonic muscle atrophy to lower sonic muscle tensions and resonant frequencies.

The swimbladder properties determine the shape of the oscillations occurring after the first two half-cycles in the SSMT model waveform. The frequency content of these later oscillations depends on the swimbladder resonant frequency and would be less sensitive to seasonal sonic muscle changes. Hence, the later oscillations in the weakfish purr waveform could be a better indicator of fish size, as swimbladder resonant frequency varies with fish size (Løvik and Hovem, 1979).

C. Sound production in other species

Many other species, including other Sciaenids, the oyster toadfish, the northern midshipman, and the northern searobin, produce sound using sonic muscles to drive the swimbladder. The SSMT model may lead to better understanding of their sound-producing mechanisms.

The spotted seatrout, a congener to the weakfish with virtually the same swimbladder and sonic muscles, produces different waveforms from those of the weakfish (Mok and Gilmore, 1983; Sprague *et al.*, 2000). The spotted seatrout waveforms must be the result of a different excitation of the sonic muscles than those causing weakfish waveforms. A model of spotted seatrout sound production will be the subject of a future study.

The silver perch is a Sciaenid that produces a higher frequency sound than the weakfish (Mok and Gilmore, 1983; Luczkovich *et al.*, 1999a, 2000; Sprague *et al.*, 2000). Examination of the waveform of the silver perch ''knock'' suggests that the silver perch sonic muscle may have a twitch that is even faster than that of the weakfish.

VI. CONCLUSIONS

The SSMT model reproduces the waveform of weakfish purr with a sound-producing mechanism requiring a single sonic muscle twitch. The model is based on the theories of driven damped oscillation and radiation of sound by a vibrating sphere. The parameters used for the model are within the range of experimentally measured values, suggesting that the model may be interpreted realistically. The SSMT model may lead to a better understanding of the sound-production mechanisms in weakfish and other Sciaenids as well as in other fishes that use their swimbladders and associated muscles to produce sound.

ACKNOWLEDGMENTS

This work would not have been possible without the help of others. Thanks to Joseph J. Luczkovich and Hal J. Daniel for their helpful discussions and suggestions. Thanks to Martin A. Connaughton and Stephen F. Norton for helping me to better understand muscle twitches and sonic muscle anatomy.

Bass, A. H., and Marchaterre, M. A. (**1989**). ''Sound-generating (sonic) motor system in a Teleost fish (*Porichthys notatus*): Sexual polymorphism in the ultrastructure of myofibrils,'' J. Comp. Neurol. **286**, 141–153.

Cohen, M. J., and Winn, H. E. (**1967**). ''Electrophysical observations on hearing and sound production in the fish *Porichthys notatus*,'' J. Exp. Zool. **165**, 355–370.

Connaughton, M. A., Taylor, M. H., and Fine, M. L. (**2000**). ''Effects of fish size and temperature on weakfish disturbance calls: Implications for the mechanism of sound generation,'' J. Exp. Biol. **203**(9), 1503–1512.

Connaughton, M. A., Fine, M. L., and Taylor, M. H. (**1997**). ''The effects of seasonal hypertrophy and atrophy on fiber morphology, metabolic substrate concentration and sound characteristics of the weakfish sonic muscle,'' J. Env. Biol. **200**(18), 2449–2457.

Connaughton, M. A., and Taylor, M. H. (**1996**). ''Drumming, courtship, and spawning behavior in captive weakfish, *Cynoscion regalis*,'' Copeia **1996**(1), 195–199.

Connaughton, M. A., and Taylor, M. H. (**1995**). ''Seasonal and daily cycles in sound production associated with spawning in weakfish, *Cynoscion regalis*,'' Environ. Biol. Fish. **42**(3), 233–240.

Crawford, J. D., Hagedorn, M., and Hopkins, C. D. (**1986**). ''Acoustic communication in an electric fish, *Pollimyrus isidori* (Mormyridae),'' J. Comp. Physiol. A **159**, 297–310.

Demski, L. S., Gerald, J. W., and Popper, A. N. (**1973**). ''Central and peripheral mechanisms in teloest sound production,'' Am. Zool. **13**, 1141–1167.

Fine, M. L., Winn, H. E., and Olla, B. L. (**1977**). ''Communication in fishes,'' in *How Animals Communicate*, edited by T. A. Sebeok (Indiana University Press, Bloomington, IN), pp. 472–518.

Fine, M. L. (**1997**). ''Endocrinology of sound production in fishes,'' Mar. Freshwater Behav. Physiol. **29**, 23–45.

Fish, M. P., and Mowbray, W. H. (**1970**). *Sounds of the Western North Atlantic Fishes* (Johns Hopkins, Baltimore, MD).

Guest, W. C., and Lasswell, J. L. (**1978**). ''A note on courtship behavior and sound production of red drum,'' Copeia **1978**(2), 337–338.

Harris, G. G. (**1964**). ''Considerations on the physics of sound production by fishes,'' in *Marine Bio-Acoustics*, edited by W. N. Tavolga (Macmillan, New York), pp. 233–247.

Hill, G. L., Fine, M. L., and Musik, J. A. (**1987**). ''Ontogeny of the sexually dimorphic sonic muscles in three Sciaenid species,'' Copeia **1987**(3), 708–713.

Løvik, A., and Hovem, J. M. (**1979**). ''An experimental investigation of swimbladder resonance in fishes,'' J. Acoust. Soc. Am. **66**(3), 850–854.

Luczkovich, J. J., Daniel III, H. J., Hutchinson, M., Jenkins, T., Johnson, S. E., Pullinger, R. C., and Sprague, M. W. (**2000**). ''Sounds of sex and death in the sea: Bottlenose dolphin whistles suppress mating choruses of silver perch,'' Bioacoustics **10**(4), 323–334.

Luczkovich, J. J., Sprague, M. W., Johnson, S. E., and Pullinger, R. C. (**1999a**). ''Delimiting spawning areas of weakfish, *Cynoscion regalis* (family Sciaenidae) in Pamlico Sound, North Carolina using passive hydroacoustic surveys,'' Bioacoustics **10**, 143–160.

Luczkovich, J. J., Daniel III, H. J., Sprague, M. W., Johnson, S. E., Pullinger, R. C., Jenkins, T., and Hutchinson, M. (**1999b**). ''Characterization of critical spawning habitats of weakfish, spotted seatrout and red drum in Pamlico Sound using hydrophone surveys'' (North Carolina Dept. of Environ. and Nat. Resour., Div. Mar. Fish., Morehead City, NC).

Mann, D. A., Bowers-Altman, J., and Rountree, R. A. (**1997**). ''Sounds produced by the striped cusk-eel *Ophidion marginatum* (Ophidiidae) during courtship and spawning,'' Copeia **1997**(3), 610–612.

Marion, J. B., and Thornton, S. T. (**1995**). *Classical Dynamics of Particles and Systems* (Saunders College Publishing, Fort Worth, TX).

Mok, H. K., and Gilmore, R. G. (**1983**). ''Analysis of sound production in estuarine fish aggregations of *Pogonias cromis, Bairdiella chrysoura*, and *Cynoscion nebulosus* (Sciaenidae),'' Bull. Inst. Zool. Acad. Sinica **22**(2), 157–186.

Myrberg, A. A. (**1981**). ''Sound communication and interception in fishes,'' in *Hearing and Sound Communication in Fishes*, edited by W. N. Tavolga, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 395–425.
Myrberg, A. A., Kramer, E., and Heinecke, P. (**1965**). ''Sound production by cichlid fishes,'' Science **149**, 555–558.
Ono, R. D., and Poss, S. G. (**1982**). ''Structure and innervation of the swim bladder musculature in the weakfish, *Cynoscion regalis* (Teliostei: Sciaenidae),'' Can. J. Zool. **60**, 1955–1967.
Pierce, A. D. (**1989**). *Acoustics An Introduction to its Physical Principles and Applications* (Acoustical Society of America, Woodbury, NY).
Rome, L. C., Syme, D. A., Hollingsworth, S., and Lindstedt, S. L. (**1996**). ''The whistle and the rattle: The design of sound producing muscles,'' Proc. Natl. Acad. Sci. U.S.A. **93**, 8095–8100.
Skoglund, C. R. (**1961**). ''Functional analysis of swimbladder muscles engaged in sound production of the toadfish,'' J. Biophys. Biochem. Cytol. **10** (Suppl.), 187–200.
Smith, H. M. (**1905**). ''The drumming of the drum-fishes (Sciaenidae),'' Science **22**, 376–378.
Sprague, M. W., Luczkovich, J. J., Pullinger, R. C., Johnson, S. E., Jenkins, T., and Daniel, III, H. J. (**2000**). ''Using spectral analysis to identify drumming sounds of some North Carolina fishes in the family *Sciaenidae*,'' J. Elish. Mitch. Sci. Soc. **116**(2), 124–145.
Tavolga, W. N. (**1964**). ''Sonic characteristics and mechanisms in marine fishes,'' in *Marine Bio-Acoustics*, edited by W. N. Tavolga (Macmillan, New York), pp. 195–211.
Tower, R. W. (**1908**). ''The production of sound in the drumfishes, the sea-robin and the toadfish,'' Ann. (N.Y.) Acad. Sci. **18**, 149–180.
Winn, H. E. (**1964**). ''The biological significance of fish sounds,'' in *Marine Bio-Acoustics*, edited by W. N. Tavolga (Macmillan, New York), pp. 213–231.
Wolfram, S. (**1996**). *The Mathematica Book*, 3rd ed. (Wolfram Media/Cambridge University Press, Champaign, IL).

LETTERS TO THE EDITOR

This Letters section is for publishing (a) brief acoustical research or applied acoustical reports, (b) comments on articles or letters previously published in this Journal, and (c) a reply by the article author to criticism by the Letter author in (b). Extensive reports should be submitted as articles, not in a letter series. Letters are peer-reviewed on the same basis as articles, but usually require less review time before acceptance. Letters cannot exceed four printed pages (approximately 3000–4000 words) including figures, tables, references, and a required abstract of about 100 words.

Variations in temporal patterns of speech production among speakers of English

Bruce L. Smith[a)]
Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois 60208

(Received 17 April 2000; accepted for publication 12 July 2000)

In many languages, some segments are longer/shorter than others due to phonetic context, position within an utterance, etc. However, results from most such studies have typically considered group averages. The present study examined productions by ten subjects to assess the extent to which they individually showed various temporal patterns viewed as characteristic of English. While all subjects demonstrated most of the patterns considered, there were instances in which one or more of them did not. Even when they did manifest expected patterns, there was typically quite a large range of performance across speakers. © *2000 Acoustical Society of America.*
[S0001-4966(00)04010-8]

PACS numbers: 43.70.Fq [AL]

I. INTRODUCTION

Many studies have determined that in English and other languages, some segments tend to be longer or shorter than others as a consequence of factors such as phonetic context, position within an utterance, and due to intrinsic characteristics of the segments themselves (e.g., Peterson and Lehiste, 1960; Lehiste, 1970; Klatt, 1976; Oller, 1973; de Jong, 1991). For instance, English is commonly described as showing patterns such as phrase-final vowel lengthening, vowel lengthening before voiced obstruents, etc. However, results from most studies providing such descriptions have typically considered tendencies involving group averages. Although various trends are commonly observed when investigating group findings, comparatively little is known about the extent to which *individual* subjects exhibit these different temporal patterns and what kind of variation there is across speakers.

Although variability across speakers has been considered in some investigations (e.g., van den Heuvel, Rietveld and Cranen, 1994), it is often largely overlooked because of interest in gaining insights about *general* patterns. Even when the performance of individual subjects has been noted in certain studies, there has often been limited evaluation of such data. In other instances, relatively few subjects have been utilized, which eliminates the possibility of systematically examining range of performance across speakers (e.g., Chen, 1970; Lehiste, 1972). Although some amount of variability among subjects is expected, the extent to which individual speakers manifest different temporal patterns said to be characteristic of English could have significant implications for areas such as first and second language learning, speech recognition, etc. It is commonly assumed, for instance, that children learning English as their native language, as well as those acquiring English as a second language, must learn various language-specific temporal patterns (e.g., Smith, 1978; de Jong and Zawaydeh, 1999). However, to the extent there is substantial variability among native adult speakers for different temporal parameters, it is possible that the need for children or non-native speakers to learn specific patterns considered representative of English could be less critical than might be expected.

An important issue regarding variation across subjects concerns what the cause of different temporal patterns is. It is typically suggested that most of them are either a result of ''universal'' aspects of speech production or that they are learned for language-specific reasons. It is also possible that certain tendencies reflect intrinsic characteristics of speech that are enhanced for phonological purposes. Whether such patterns are learned or are inherently a consequence of speech production/perception generally remains rather unclear. This may be at least partly because of the implicit assumption that all speakers demonstrate reasonably similar performance regarding such patterns. It is possible, however, that analyzing the speech of individual subjects might lead to a clearer understanding of the nature of such temporal characteristics. For example, temporal parameters intrinsic to the speech production process might show less variability across speakers than patterns that are learned.

While it is essential to have an understanding of speech

a)Electronic mail: b-smith2@northwestern.edu

production patterns in a language based on averaged data from groups of subjects, it is also important to consider the extent to which individual speakers manifest variations in temporal patterns. Because data from several studies (e.g., Oller, 1973) have indicated that at least some temporal properties of English show a certain amount of intersubject variability, this issue was examined in detail in the present investigation. It was hypothesized that there would be considerable variation across subjects for the temporal patterns examined, but that some measures would show less variation than others, potentially as a function of whether they involved language-specific learning or were more ''universal'' in nature. Given the focus on individual subjects, some other factors that have not been considered in group-oriented studies were also examined. Specifically, because subjects' speaking rates can differ (Tsao and Weismer, 1997) and because certain temporal patterns can vary as a function of speaking rate (Miller, Grosjean, and Lomanto, 1984), several temporal characteristics were evaluated relative to a measure of speakers' ''nontarget'' segments. Also, potential inter-relationships among various temporal patterns were evaluated to determine whether speakers who show one temporal property to a greater extent than other subjects might manifest a ''stronger'' occurrence of other temporal patterns (Klatt, 1973).

II. METHODS

A. Subjects and stimuli. Ten native adult monolingual speakers of English (seven females, three males) ranging from 19–41 years of age ($M=27$ years) were recorded. By self-report and informal observation, all subjects were determined to be free of any speech or hearing problems. Four subjects were raised in Illinois, two in California, one in Iowa, one in Ohio, one in Pennsylvania, and one in New York. Each subject produced 15 randomized repetitions of each of the 11 stimuli: coat, goat, back, bag, cap, cab, keep, peak, pack, package, and backpack, spoken in the carrier phrase, ''I like to say —— every week.'' Fifteen randomized repetitions of these same words were also produced in the phrase, ''Now I am going to say ——.'' Whether the stimuli were produced first in the phrase involving final or nonfinal position was randomly determined. For each carrier phrase, subjects read the 11 stimuli from three different randomized lists, cycling through them five times each. Two extra words that were not analyzed were included at the beginning and end of each of the lists to avoid possible ''end-of-list'' effects. Subjects were asked to speak naturally as they produced the utterances, but no attempt was made to control how slowly or quickly they spoke. Thus, in addition to potential effects of different dialects, it is also possible that some of the variation reported below was due to certain subjects having spoken in a more ''conversational'' style versus others who may have attempted to speak particularly clearly. In all cases, however, the subjects' productions were judged to be acceptable.

B. Procedures. Each speaker's productions were recorded in a sound booth and digitized at 22 kHz using SOUNDEDIT 16 (2.0.7) software on a Power Macintosh 7100 computer. Utilizing digital oscillograms displayed on a Power Macintosh 8500, the following temporal patterns were examined: (a) vowel and consonant lengthening in utterance-final versus nonfinal position; (b) vowel lengthening before voiced versus voiceless stops; (c) lengthening of voiceless stop closure intervals relative to voiced stops; (d) lengthening of low vowels relative to high vowels; and (e) lengthening of vowels in monosyllables relative to disyllables. Segmentation of the various consonant and vowel intervals involved in each of these measures was based on commonly utilized acoustic characteristics associated with substantial changes in waveform shape and/or amplitude, consonant release bursts, etc. Interjudge reliability averaged approximately 12 ms (8%) for vowels and 12 ms (11%) for consonants.

III. RESULTS

Results for the various temporal characteristics examined are presented in Fig. 1. The left half of the figure shows data for the ten individual subjects for each temporal measure considered; the right side presents group data, shown as boxplots, for these same parameters. For the boxplots, the vertical line through a box is the median. The two ends of a box show the 25th and 75th percentiles, and the 10th and 90th percentiles are shown by the lines outside the box. The circles show the most extreme individual values beyond the 10th and 90th percentiles. On both sides of Fig. 1, values greater than 1.00 (i.e., to the right of the dotted vertical lines) indicate that lengthening occurred, whereas values of 1.00 or less indicate no lengthening occurred.

Figure 1 indicates that for *group* averages, the ''expected'' temporal patterns were observed in all cases. Although there were instances in which certain subjects did not manifest patterns viewed as characteristic of English, most *individual* speakers also showed these patterns to at least some extent. The right half of Fig. 1 in the top row shows, for example, that when averaging across the stimuli for evaluating phrase-final vowel lengthening (i.e., comparing the vowels from ''back, bag, cab, cap, keep, peak, and pack'' in the final versus nonfinal carrier phrases), phrase-final vowels were approximately 23% longer than nonfinal vowels for the group ($t=3.80$; $df=9$; $p<0.01$). The left-hand portion of Fig. 1 indicates, however, that although nine of the ten speakers showed phrase-final vowel lengthening to at least some degree, subject Rb did not have any final-syllable lengthening. Her final/nonfinal ratio was 0.84, indicating that her *nonfinal* vowels were about 16% longer than her final vowels. (Rb did not show phrase-final lengthening for any of the seven individual stimuli; her values ranged from 0.73 for cap to 0.99 for bag.) Among the nine subjects who did show final vowel lengthening, averages ranged from a final/nonfinal ratio of 1.05 for C to 1.57 for M. Even among these subjects whose final vowels averaged longer than their nonfinal vowels, however, two of them (C and E) showed either shortening or no lengthening of phrase-final vowels for one or two of the seven individual words.

Among the parameters considered, phrase-final vowel lengthening might be viewed as the most extreme departure from what was expected (i.e., since Rb showed phrase-final vowel shortening). Even if this example is somewhat of an

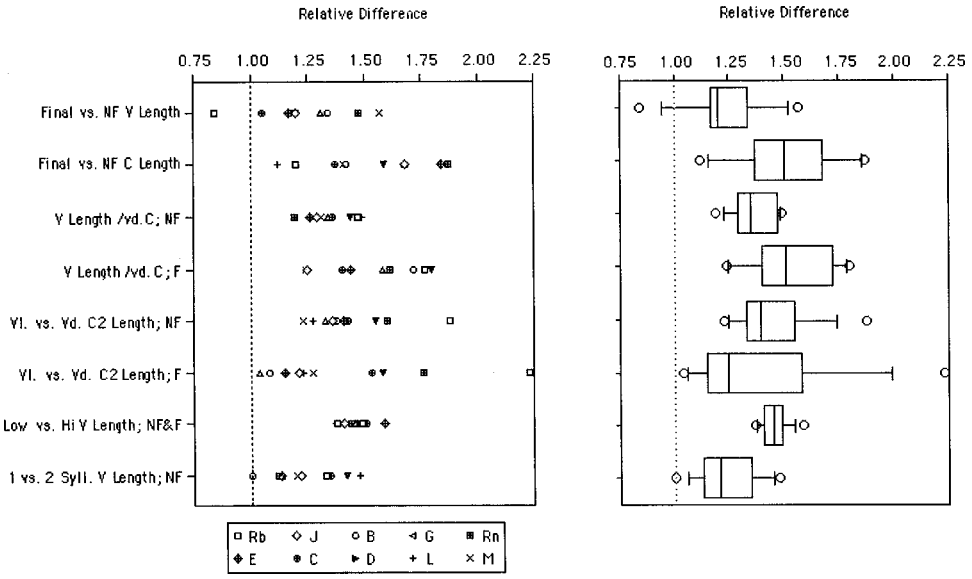


FIG. 1. Summary plot of the ten individual subjects' various temporal measures (left) and a boxplot summary of group results (right). As shown from top to bottom along the left-hand side of the figure, the parameters examined were: vowel (V) lengthening in phrase-final versus nonfinal (NF) position, consonant (C) lengthening in phrase-final versus nonfinal position, vowel lengthening before voiced consonants in nonfinal syllables, vowel lengthening before voiced consonants (vd) in phrase-final (F) syllables, lengthening of voiceless (Vl) relative to voiced stop occlusion in postvocalic (C2) position in nonfinal syllables, lengthening of voiceless versus voiced stop occlusion in postvocalic position in phrase-final syllables, low versus high vowel duration in nonfinal and final positions combined, and vowel duration in monosyllabic words as compared to disyllabic words in nonfinal position.

anomaly, however, it can be seen in the left side of Fig. 1 that several other measures also showed quite large ranges of performance among the ten subjects, including some with values at or near 1.00. For instance, the bottom row on the right side of Fig. 1 indicates that for the group as a whole, stressed vowels of monosyllables in the nonfinal carrier phrase (''I like to say —— every week'') averaged approximately 23% longer ($t=5.14$; $df=9$; $p<0.001$) than stressed vowels of disyllables (i.e., pack versus package and back versus backpack) from that phrase. As shown on the left side of the figure for this measure, however, values among the individual subjects ranged from no difference between vowels in monosyllables versus disyllables (B) to approximately 50% longer vowels (L) in monosyllables. It is also worth noting in the left half of Fig. 1 that no subject consistently showed extreme values across all the patterns investigated. For instance, several different speakers exhibited values at or near 1.00 for the various temporal parameters, and different subjects can also be seen at the ''extreme lengthening'' end of the ranges. Rb is the only subject who was somewhat of an ''outlier'' for several measures.

Figure 1 also indicates that values for the individual subjects were generally quite evenly distributed across the ranges for each of the temporal parameters. Other than the distribution for postvocalic ($C2$) voiceless versus voiced consonant closure duration in the final-position context (the third measure from the bottom), there were few instances of highly skewed distributions. Moreover, had additional speakers been studied, it is likely that the ranges in Fig. 1 would not have increased substantially; rather, additional subjects probably would have primarily filled in ''gaps'' that may exist because only ten subjects were studied. However, if including more subjects did extend the ranges of any parameters significantly, this would only provide additional support for the observation that individual speakers often vary considerably in producing temporal patterns viewed as characteristic of English. Although differences among subjects for the most extreme values of each parameter were generally quite large, one pattern that exhibited quite limited variation across speakers was the tendency for low vowels to be longer than high vowels ($t=22.25$; $df=9$; $p<0.0001$; for cap versus keep and pack versus peak for both the final and nonfinal phrases). As shown in the second to last row on the left side of Fig. 1, not only did each subject show this

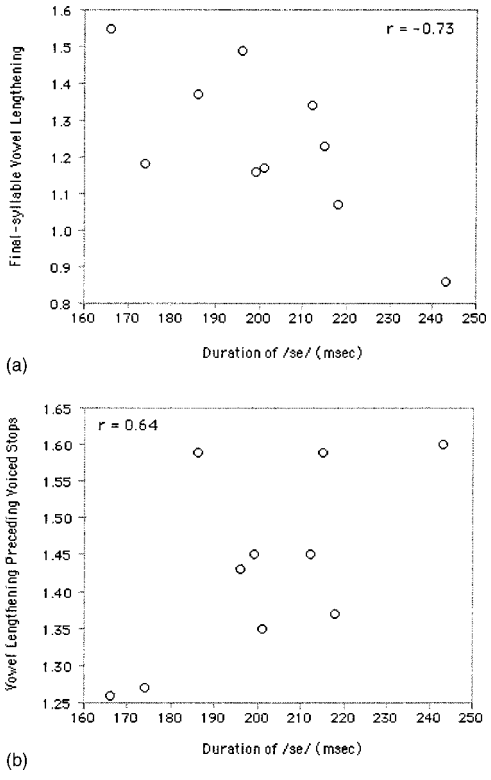


FIG. 2. (a) Scatterplot for ten subjects' durations of ''say'' from the nonfinal carrier phrase and their final-syllable vowel lengthening ratio. (b) Scatterplot for ten subjects' durations of say from the nonfinal carrier phrase and their ratio of vowel lengthening before voiced versus voiceless stops.

pattern, but their individual values ranged between about 37% and 59% greater for low vowels; this 22% difference is only about one-half to one-third of the ranges of most of the other parameters, e.g., compare utterance-final versus nonfinal consonant length (row 2).

Another factor considered was the extent to which individual subjects' ''nontarget'' durations varied and how this might relate to the specific temporal patterns examined. To assess this, the syllable /se/ from ''I like to *say* —— every week'' was measured; for each subject, the average duration of ''say'' was based on 165 productions (i.e., 15 repetitions of the 11 target words produced in this phrase). The average duration of say for the group was 201 ms, ranging from 166 ms for M to 243 ms for Rb. Correlations were computed for nonfinal say duration relative to several temporal patterns. For example, as shown in Fig. 2(a), phrase-final vowel lengthening was negatively correlated with the duration of say ($r=-0.73$; $df=8$; $p<0.05$, two-tailed). That is, subjects with longer average durations for say tended to show less final vowel lengthening, while subjects with shorter /se/ durations had, in relative terms, greater final-vowel lengthening. In contrast, Fig. 2(b) indicates that vowel lengthening preceding voiced versus voiceless stops had a positive correlation with the duration of say ($r=0.64$; $df=8$; $p<0.05$, two-tailed). Thus, subjects with longer say durations tended to manifest greater vowel lengthening before voiced versus voiceless stops, while subjects with shorter say durations showed less vowel lengthening before voiced versus voiceless stops. Because the order of subjects along the *x* axis is the same in Figs. 2(a) and (b), it can be seen that subjects with shorter average durations for say had the greatest amount of final-syllable vowel lengthening but the least vowel lengthening before voiced versus voiceless stops; in contrast, subjects with longer durations for say showed lesser amounts of final-syllable vowel lengthening but more vowel lengthening before voiced stops.

IV. DISCUSSION

On average, the temporal parameters examined showed the same group-based patterns observed in previous investigations (e.g., Peterson and Lehiste, 1960; Oller, 1973; de Jong, 1991). However, there was often considerable variation among the ten subjects in the extent to which they showed these patterns. Although most subjects generally exhibited the expected patterns to at least some extent, there were instances in which one or more of them did not—if not on average, then for certain individual stimuli. Even among subjects who did show the expected patterns, however, the range of performance was often quite large. Although some amount of variation had been expected among the speakers, the ranges of performance were greater than might have been anticipated, considering what is commonly reported to be characteristic of English. This variation among individual subjects has some potentially important ramifications regarding young children learning English as their native language and non-native speakers acquiring English as a second language. That is, they may be exposed to more temporal variation than has previously been realized. Although native adult speakers of English are quite skilled in ''normalizing'' variations to which they are exposed (Ohala, 1993), less-skilled first- or second-language learners might experience a certain amount of confusion as to what they must learn to produce in acquiring English temporal patterns (Ratner, 1986). Alternatively, it could be that the specific patterns these learners must acquire are not as critical as might be assumed.

Another factor that appeared to relate to the range of performance across speakers was differences in speaking rate. Subjects with the longest or shortest syllable durations tended to show the most extreme, and sometimes opposite, temporal patterns. Thus, subjects with greater average syllable durations tended to produce larger relative amounts of vowel lengthening before voiced versus voiceless obstruents than subjects with shorter syllable durations; in contrast, subjects with longer average syllable durations tended to show less (relative) vowel lengthening in phrase-final position than subjects with shorter syllable durations. It is not clear why two temporal parameters that are both manifestations of conditioned ''vowel lengthening'' would show opposite patterns relative to speakers' overall syllable durations. It could be that phrase-final vowel lengthening is influenced more by ''global'' timing factors related to various prosodic properties, whereas vowel lengthening before voiced stops is affected more by the ''local,'' phonetic context.

In conclusion, it appears that although English, in general, exhibits various temporal patterns related to phonetic context, position-in-utterance, etc., a reasonably large range of performance can be observed among individual subjects in the extent to which they manifest such patterns. Not all talkers necessarily show all temporal properties commonly considered characteristic of English. Though most speakers do manifest these temporal properties to at least some degree, the range across subjects can be quite large for many measures. If it is reasonable to assume that temporal parameters inherently due to the speech production process itself tend to show less variability across speakers than patterns that are learned, this would suggest that the duration contrast between high versus low vowels was the pattern in the present study that least likely involved language-specific learning. In contrast, patterns associated with various aspects of final vowel or consonant length were typically some of the most variable across subjects and may, therefore, be affected more easily by factors such as speaking rate or other prosodic influences that may be learned and, therefore, more speaker dependent. Ultimately, the extent to which variation across speakers is a result of learned and/or inherent influences largely remains unclear, and it is likely that the causes of such ''patterns'' are more complex than often assumed.

Chen, M. (**1970**). ''Vowel length variation as a function of the voicing of the consonant environment,'' Phonetica **22**, 129–159.

de Jong, K. (**1991**). ''An articulatory study of consonant-induced vowel duration changes in English,'' Phonetica **48**, 1–17.

de Jong, K., and Zawaydeh, B. A. (**1999**). ''Stress, duration, and intonation in Arabic word-level prosody,'' J. Phonetics **27**, 3–22.

Klatt, D. H. (**1973**). ''Interaction between two factors that influence vowel duration,'' J. Acoust. Soc. Am. **54**, 1102–1104.

Klatt, D. H. (**1976**). ''Linguistic uses of segmental duration in English: Acoustic and perceptual evidence,'' J. Acoust. Soc. Am. **59**, 1208–1221.

Lehiste, I. (**1970**). *Suprasegmentals* (M.I.T. Press, Cambridge, MA).

Lehiste, I. (**1972**). ''Timing of utterances and linguistic boundaries,'' J. Acoust. Soc. Am. **51**, 2018–2024.
Miller, J. L., Grosjean, F., and Lomanto, C. (**1984**). ''Articulation rate and its variability in spontaneous speech: A reanalysis and some implications,'' Phonetica **41**, 215–225.
Ohala, J. J. (**1993**). ''Coarticulation and phonology,'' Lang. Speech **36**, 155–170.
Oller, D. K. (**1973**). ''The effect of position in utterance on speech segment duration in English,'' J. Acoust. Soc. Am. **54**, 1235–1247.
Peterson, G. E., and Lehiste, I. (**1960**). ''Duration of syllable nuclei in English,'' J. Acoust. Soc. Am. **32**, 693–703.
Ratner, N. B. (**1986**). ''Durational cues which mark clause boundaries in mother–child speech,'' J. Phonetics **14**, 303–309.
Smith, B. L. (**1978**). ''Temporal aspects of English speech production: A developmental perspective,'' J. Phonetics **6**, 37–67.
Tsao, Y-C., and Weismer, G. (**1997**). ''Interspeaker variation in habitual speaking rate: Evidence for a neuromuscular component,'' J. Speech, Lang., Hear. Res. **40**, 858–866.
van den Heuvel, H., Rietveld, T., and Cranen, B. (**1994**). ''Methodological aspects of segment- and speaker-related variability. A study of segmental durations in Dutch,'' J. Phonetics **22**, 389–406.

Erratum: "Power transmission from a vibrating body to a circular cylindrical shell through passive and active isolators" [J. Acoust. Soc. Am. 101, 1479–1491 (1997)]

Carl Q. Howard and Colin H. Hansen
Department of Mechanical Engineering, University of Adelaide, S.A. 5005, Australia

(Received 16 August 1999; accepted for publication 2 June 2000)

[S0001-4966(00)03010-1]

I. INTRODUCTION

This erratum presents the corrections to the mathematical model by Howard *et al.*[1] that describes the power transmission into a circular cylindrical shell through either passive or active vibration isolators. The model can be checked for consistency by calculating the power transmission into the cylinder for passive vibration isolation and then rotating the coordinate system by 90 degrees and checking that the same values of power transmission are obtained. When this check was performed it was found that different values of power transmission were obtained for two different orientations of the coordinate system. This check is satisfied for the corrected model described in this erratum.

A simply supported closed circular cylindrical shell has two sets of harmonic solutions that are orthogonal to each other. This property does not occur for beam or plate structures and makes the analysis of closed cylindrical structures slightly more complicated than beam or plate systems. The first set is termed the ''odd'' modes because the radial vibration $w(s,\theta)$ is described by a sine function in the circumferential direction and the sine function is asymmetric about the origin. The second set is termed ''even'' modes because a cosine function is used which is symmetric about the origin. The displacement of the cylinder is given by the summation of these even and odd modes. For vibration problems which analyze the displacement of a closed cylindrical shell, the contribution of these modes can be considered separately and then added together. In the investigation of power transmission into a cylinder with an attached structure, the even and odd modes need to be considered at the same time because the calculation of power involves taking the real part of the product of force and velocity. For example, say $f_1^r+jf_1^i$ and $f_2^r+jf_2^i$ represent the complex force driving the cylinder for the even and odd modes and $v_1^r+jv_1^i$ and $v_2^r+jv_2^i$ represent the complex velocity response of the cylinder for the even and odd modes, where $f_1^r,\ldots,v_2^i$ are real numbers. As these two responses need to be added together to calculate the power transmissions, it follows that

$$\begin{aligned}&\mathrm{Re}\{\text{Force}\times\text{Velocity}\}\\&=\mathrm{Re}\{((f_1^r+jf_1^i)+(f_2^r+jf_2^i))\times((v_1^r+jv_1^i)+(v_2^r+jv_2^i))\}\\&\neq\mathrm{Re}\{(f_1^r+jf_1^i)\times(v_1^r+jv_1^i)\}+\mathrm{Re}\{(f_2^r+jf_2^i)\times(v_2^r+jv_2^i)\}.\end{aligned}$$

This means that the total response of the cylinder has to be calculated first, by the summation of the odd and even modal responses, before the power is calculated by taking the real part of the product of force and velocity. The model described by Howard *et al.*[1] calculated the power transmission for the odd and even modal responses separately and then added them together, which is incorrect. The correct method is to calculate the sum of the odd and even modal responses and then calculate the power transmission.

II. CORRECTED EQUATIONS

The correct equations are described below. References are made to the equation numbers in Howard *et al.*[1] and the replacement equation is described.

The displacements of the cylindrical shell along the three axes, as given by the contribution of the even and odd modes described in Eqs. (21a)–(21c), are incorrect. The correct equations are

$$\xi_s(s,\theta)=\sum_{m=1,n=0}^{\infty}a_{m,n}\cos\lambda s\sin n\theta+a'_{m,n}\cos\lambda s\cos n\theta,$$

$$\xi_\theta(s,\theta)=\sum_{m=1,n=0}^{\infty}b_{m,n}\sin\lambda s\cos n\theta+b'_{m,n}\sin\lambda s\sin n\theta,$$

$$w(s,\theta)=\sum_{m=1,n=0}^{\infty}c_{m,n}\sin\lambda s\sin n\theta+c'_{m,n}\sin\lambda s\cos n\theta,$$

where $a_{m,n},a'_{m,n},\ldots,c_{m,n},c'_{m,n}$ are modal amplitude constants for mode (m,n), L_0 is the length of the cylinder, and $\lambda=m\pi R/L_0$.

The uncoupled shell characteristic matrix $\mathbf{Z}_s$ shown in Eq. (31) is incorrect. The correct matrix must include the odd and even modal responses given by

$$\mathbf{Z}_s=\begin{bmatrix}\Omega_1 & & & & & \\ & \ddots & & & & \\ & & \Omega_P & & & \\ & & & \Omega'_1 & & \\ & & & & \ddots & \\ & & & & & \Omega'_P\end{bmatrix}-\sum_{J=1}^{L_1}m_J^b\omega^2\begin{bmatrix}\psi_k(\boldsymbol{\sigma}_J)+\psi_k(\boldsymbol{\sigma}_J) & \psi'_k(\boldsymbol{\sigma}_J)+\psi_k(\boldsymbol{\sigma}_J)\\ \psi_k(\boldsymbol{\sigma}_J)+\psi'_k(\boldsymbol{\sigma}_J) & \psi'_k(\boldsymbol{\sigma}_J)+\psi'_k(\boldsymbol{\sigma}_J)\end{bmatrix}.$$

The force location matrix $\mathbf{R}_J^b$ in Eq. (36) is incorrect and should be

 0001-4966/2000/108(5)/2443/2/$17.00

$$\mathbf{R}_J^b=\begin{bmatrix} \frac{h}{2R}\psi_{1s}(\boldsymbol{\sigma}_J) & \frac{h}{2R}\psi_{1\theta}(\boldsymbol{\sigma}_J) & \psi_1(\boldsymbol{\sigma}_J) & \frac{1}{R}\psi_{1\theta}(\boldsymbol{\sigma}_J) & -\frac{1}{R}\psi_{1s}(\boldsymbol{\sigma}_J) & 0 \\ \frac{h}{2R}\psi_{2s}(\boldsymbol{\sigma}_J) & \frac{h}{2R}\psi_{2\theta}(\boldsymbol{\sigma}_J) & \psi_2(\boldsymbol{\sigma}_J) & \frac{1}{R}\psi_{2\theta}(\boldsymbol{\sigma}_J) & -\frac{1}{R}\psi_{2s}(\boldsymbol{\sigma}_J) & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{h}{2R}\psi_{Ps}(\boldsymbol{\sigma}_J) & \frac{h}{2R}\psi_{P\theta}(\boldsymbol{\sigma}_J) & \psi_P(\boldsymbol{\sigma}_J) & \frac{1}{R}\psi_{P\theta}(\boldsymbol{\sigma}_J) & -\frac{1}{R}\psi_{Ps}(\boldsymbol{\sigma}_J) & 0 \end{bmatrix}$$
$$+\begin{bmatrix} \frac{h}{2R}\psi'_{1s}(\boldsymbol{\sigma}_J) & \frac{h}{2R}\psi'_{1\theta}(\boldsymbol{\sigma}_J) & \psi'_1(\boldsymbol{\sigma}_J) & \frac{1}{R}\psi'_{1\theta}(\boldsymbol{\sigma}_J) & -\frac{1}{R}\psi'_{1s}(\boldsymbol{\sigma}_J) & 0 \\ \frac{h}{2R}\psi'_{2s}(\boldsymbol{\sigma}_J) & \frac{h}{2R}\psi'_{2\theta}(\boldsymbol{\sigma}_J) & \psi'_2(\boldsymbol{\sigma}_J) & \frac{1}{R}\psi'_{2\theta}(\boldsymbol{\sigma}_J) & -\frac{1}{R}\psi'_{2s}(\boldsymbol{\sigma}_J) & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{h}{2R}\psi'_{Ps}(\boldsymbol{\sigma}_J) & \frac{h}{2R}\psi'_{P\theta}(\boldsymbol{\sigma}_J) & \psi'_P(\boldsymbol{\sigma}_J) & \frac{1}{R}\psi'_{P\theta}(\boldsymbol{\sigma}_J) & -\frac{1}{R}\psi'_{Ps}(\boldsymbol{\sigma}_J) & 0 \end{bmatrix},$$

and therefore the displacement of the cylindrical shell in Eq. (35) should be

$$\mathbf{W}_J=[\xi_s \quad \xi_\theta \quad w \quad \theta_s \quad \theta_\theta \quad \theta_w]^T=[\boldsymbol{\Gamma}_J \quad \boldsymbol{\Gamma}'_J]\begin{bmatrix}\mathbf{c}_k\\ \mathbf{c}'_k\end{bmatrix},$$

where Γ_J is given as follows:

$$\boldsymbol{\Gamma}_J=\begin{bmatrix} \frac{h}{2R}\psi_{1s}(\boldsymbol{\sigma}_J) & \frac{h}{2R}\psi_{2s}(\boldsymbol{\sigma}_J) & \cdots & \frac{h}{2R}\psi_{Ps}(\boldsymbol{\sigma}_J) \\ \frac{h}{2R}\psi_{1\theta}(\boldsymbol{\sigma}_J) & \frac{h}{2R}\psi_{2\theta}(\boldsymbol{\sigma}_J) & \cdots & \frac{h}{2R}\psi_{P\theta}(\boldsymbol{\sigma}_J) \\ \psi_1(\boldsymbol{\sigma}_J) & \psi_2(\boldsymbol{\sigma}_J) & \cdots & \psi_P(\boldsymbol{\sigma}_J) \\ \frac{1}{R}\psi_{1\theta}(\boldsymbol{\sigma}_J) & \frac{1}{R}\psi_{2\theta}(\boldsymbol{\sigma}_J) & \cdots & \frac{1}{R}\psi_{P\theta}(\boldsymbol{\sigma}_J) \\ -\frac{1}{R}\psi_{1s}(\boldsymbol{\sigma}_J) & -\frac{1}{R}\psi_{2s}(\boldsymbol{\sigma}_J) & \cdots & -\frac{1}{R}\psi_{Ps}(\boldsymbol{\sigma}_J) \\ 0 & 0 & \cdots & 0 \end{bmatrix},$$

and likewise for the definition of Γ' where ψ' terms are used instead of the ψ terms.

Equations (39a)–(39d) should be replaced with

$$\mathbf{A}_{11}=\mathbf{Z}_0+\sum_{J=1}^{L_1}\mathbf{R}_J^t\mathbf{K}_J[\mathbf{R}_J^t]^T,$$

$$\mathbf{A}_{12}=-\sum_{J=1}^{L_1}\mathbf{R}_J^t\mathbf{K}_J[\mathbf{T}_J]^T[\boldsymbol{\Gamma}_J\boldsymbol{\Gamma}'_J],$$

$$\mathbf{A}_{21}=-\sum_{J=1}^{L_1}\mathbf{R}_J^b\mathbf{T}_J\mathbf{K}_J[\mathbf{R}_J^t]^T,$$

$$\mathbf{A}_{22}=\mathbf{Z}_s+\sum_{J=1}^{L_1}\mathbf{R}_J^b\mathbf{T}_J\mathbf{K}_J[\mathbf{T}_J]^T[\boldsymbol{\Gamma}_J\boldsymbol{\Gamma}'_J],$$

which results in the replacement of Eqs. (51)–(54) with

$$\mathbf{G}_1=\begin{bmatrix}[\mathbf{H}_1]^H[\boldsymbol{\Gamma}_1\boldsymbol{\Gamma}'_1]^T\mathbf{T}_1 & [\mathbf{H}_1]^H[\boldsymbol{\Gamma}_2\boldsymbol{\Gamma}'_2]^T\mathbf{T}_2\\ [\mathbf{H}_2]^H[\boldsymbol{\Gamma}_1\boldsymbol{\Gamma}'_1]^T\mathbf{T}_1 & [\mathbf{H}_2]^H[\boldsymbol{\Gamma}_2\boldsymbol{\Gamma}'_2]^T\mathbf{T}_2\end{bmatrix},\quad \mathbf{G}_2=\begin{bmatrix}[\mathbf{R}_1^t]^T\mathbf{H}_3-[\mathbf{T}_1]^T[\boldsymbol{\Gamma}_1\boldsymbol{\Gamma}'_1]\mathbf{H}_1 & [\mathbf{R}_1^t]^T\mathbf{H}_4-[\mathbf{T}_1]^T[\boldsymbol{\Gamma}_1\boldsymbol{\Gamma}'_1]\mathbf{H}_2\\ [\mathbf{R}_2^t]^T\mathbf{H}_3-[\mathbf{T}_2]^T[\boldsymbol{\Gamma}_2\boldsymbol{\Gamma}'_2]\mathbf{H}_1 & [\mathbf{R}_2^t]^T\mathbf{H}_4-[\mathbf{T}_2]^T[\boldsymbol{\Gamma}_2\boldsymbol{\Gamma}'_2]\mathbf{H}_2\end{bmatrix},$$

$$\mathbf{G}_3=\begin{bmatrix}[\mathbf{R}_1^t]^T\mathbf{B}_{11}-[\mathbf{T}_1]^T[\boldsymbol{\Gamma}_1\boldsymbol{\Gamma}'_1]\mathbf{B}_{21}\\ [\mathbf{R}_2^t]^T\mathbf{B}_{11}-[\mathbf{T}_2]^T[\boldsymbol{\Gamma}_2\boldsymbol{\Gamma}'_2]\mathbf{B}_{21}\end{bmatrix}\mathbf{Q}_0,\quad \mathbf{G}_4=[\mathbf{Q}_0]^T[[\mathbf{B}_{21}]^H[\boldsymbol{\Gamma}_1\boldsymbol{\Gamma}'_1]^T\mathbf{T}_1 \quad [\mathbf{B}_{21}]^H[\boldsymbol{\Gamma}_2\boldsymbol{\Gamma}'_2]^T\mathbf{T}_2].$$

The remaining equations in the article by Howard *et al.*[1] are correct.

After recalculating the power transmission into the cylinder for the loading conditions described by Howard *et al.*,[1] the qualitative results do not change but the quantitative results increase by about 5 dB across the entire spectrum.[2]

[1] C. Q. Howard, C. H. Hansen, and J-Q. Pan, ''Power transmission from a vibrating body to a circular cylindrical shell through passive and active isolators,'' J. Acoust. Soc. Am. **101**, 1479–1491 (1997).

[2] C. Q. Howard, ''Active Isolation of Machinery Vibration from Flexible Structures,'' Ph.D. dissertation, University of Adelaide, Australia, 1999.

Session 0pNSa

NOISE-CON and Noise: Theoretical and Experimental Noise Control

Bruce E. Walker, Chair
Hersh Acoustical Engineering, 780 Lakefield Road, Unit G, Westlake Village, California 91361

Contributed Papers

1:20

0pNSa1. Case study evaluation of the Rayleigh integral method. D. W. Herrin (College of Eng., Univ. of Kentucky, 521 CRMS Bldg., Lexington, KY 40506-0108), T. W. Wu, and A. F. Seybert (Univ. of Kentucky, Lexington, KY 40506-0108)

The Rayleigh integral is based on the assumption that the vibrating surface is flat and part of an infinite rigid baffle. Though approximate, the Rayleigh integral can evaluate the sound power radiated by a vibrating surface in a fraction of the time required by the boundary element method. Three case studies were conducted to evaluate the reliability of the Rayleigh integral. Specifically, the Rayleigh integral was used to predict the sound power from a wheel-rail structure, a gearbox, and an engine. In each case, Rayleigh integral results were compared with boundary element results. The results indicated that sound power compared well with that obtained using the boundary element method. The differences were less than 5 dB except at low frequencies. However, the Rayleigh integral did not accurately predict the sound pressures at field points. The results suggested that the Rayleigh integral could be used in place of the more time intensive boundary element method for sound power calculations.

1:40

0pNSa2. Forcing function models for structures excited by low-speed flow. Y. F. Hwang and S. A. Hambric (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804)

Predictions of flow-induced flat plate vibrations using a finite-element structural model and the modified Corcos cross-spectral model as the forcing function [Hambric and Hwang, Proceedings of Inter-Noise 2000, Nice, France, 27–30 August (2000)] agree very well with the measured [Han *et al.*, J. Sound Vib. **227**(4), 685–709 (1999)] plate vibrations. However, extremely fine finite-element meshes had to be used to resolve the streamwise variations of the coherence function. This approach is not computationally tractable for most large complex structures. This paper discusses a practical approach of utilizing an approximated discrete forcing function model, which represents only the low-wavenumber region of the excitation in terms of the cross-spectra of grid forces. The high-wavenumber (convective ridge) content of the excitation is thus ignored. This method is believed more feasible in practice because the finite-element meshes now need only resolve the structural wave scale instead of the convective turbulence scale. This approximated forcing function model is assessed against the exact model and found to be satisfactory in most cases except when the structural mode wavenumbers are close to the convective ridge. [Work supported by ONR, Code 333.]

2:00

0pNSa3. A computational approach for flow-acoustic coupling in deep cavities. Paul M. Radavich, Ahmet Selamet (Dept. of Mech. Eng., The Ohio State Univ., Columbus, OH 43210), and James M. Novak (Ford Motor Co., Dearborn, MI 48121)

The quarter wave resonator, which produces a narrow band of high acoustic attenuation at regularly spaced frequency intervals, is a common type of silencer used in ducting systems. The presence of mean flow, however, is likely to promote an interaction between these acoustic resonances and the flow. The coupling for some discrete flow conditions leads to the production of both large wave amplitudes in the side branch and high noise levels in the main duct. Thus, the quarter wave resonator may become a noise source rather than a silencer. The present approach uses computational fluid dynamics to determine both the mean flow and acoustic fields simultaneously by solving the unsteady, turbulent, and compressible Navier–Stokes equations. By coupling the interaction between the mean flow fluctuations and vortices with the acoustic field, this method is capable of determining when flow-acoustic coupling occurs and how variations in the geometry and flow conditions influence the acoustic pressure amplitudes.

2:20

0pNSa4. Wall pressure fluctuations due to boundary-layer transition. Gerald C. Lauchle (Grad. Prog. in Acoust., and Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804, gcl1@psu.edu) and Sewon Park (Penn State Univ., State College, PA 16804)

Boundary-layer transition is a contributor to acoustic radiation and structural excitation. The pressure fluctuations caused by this spatially bounded, and intermittent, phenomenon encompass a very wide range of spatial wave numbers and temporal frequencies. Here, we analyze the wave vector-frequency spectrum of the wall pressure fluctuations due to subsonic boundary-layer transition as it occurs on a flat plate under zero pressure gradient conditions. Based on previous measurements of the statistics of the boundary-layer intermittency, it is found that transition induces higher low-streamwise wave number wall pressure levels than does a fully developed turbulent boundary layer (TBL) that might superficially exist at the same location and at the same Reynolds number. The transition zone span-wise wave number pressure components are virtually unchanged from the TBL case. The results suggest that transition may be more effective than the TBL in forcing structural excitation at low Mach numbers, and it may have a more intense radiated noise contribution on a per unit area basis.

2:40

0pNSa5. Hybrid active/passive control of sound radiation from panels with constrained layer damping and model predictive feedback control. Randolph Cabell (Structural Acoust. Branch, M/S 463, NASA Langley Res. Ctr., Hampton, VA 23681)

Feedback control of aircraft sidewall panels has been studied for reducing interior noise due to turbulent boundary layer excitation of the fuselage. Online adaptation of the feedback controller parameters can be used to track pressure, temperature, and structural variations, but the associated computational burden can be overwhelming. This work describes a hybrid active/passive control approach where the passive components reduce the complexity of the active system. Constrained layer damping provides damping at high frequencies, and a generalized predictive controller is used at low frequencies. Experiments were conducted on a panel subjected to broadband speaker excitation in a transmission loss facility. A piezoelectric actuator provided control input to the panel, accelerometers provided error feedback. Two sensing configurations were studied: one used full-state feedback to control radiated sound power, estimated from 15 accelerometers on the panel; the other used dynamic output feedback to control the summed responses of 4 accelerometers. Active noise reduction

for different locations of constrained layer damping are discussed. The addition of the constrained layer damping makes it possible to achieve good noise reduction with the simpler 4-accelerometer sensing configuration.

3:00

0pNSa6. Incorporation of sliding mode control into a vibration and noise suppression scheme for a rigid rotor. Scott Lewis (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804, as1103@psu.edu)

This paper presents a sliding mode control algorithm to suppress the vibration of a rigid rotor supported by magnetic bearings. It is assumed that unbalance exists in the shaft and that the two planes of the shaft motion are coupled by gyroscopic forces. It is assumed that the angular rotation is small in both the normal and transverse directions and that the force and current are related linearly. The sliding mode control law is designed to be robust to rotor unbalance and transient disturbances. A boundary layer is introduced around each sliding hyperplane to eliminate the chattering phenomenon. The results from numerical simulations are presented, which not only corroborate the validity of the proposed controller, but also show the effects of various control parameters as a function of the angular speed of the rotor. Special attention is addressed to the orbital plots of the ends of the rotor. Experimental results are compared to those obtained analytically and demonstrate how the current required by the bearings is affected by the angular velocity of the rotor. [Work sponsored by ONR.]

3:20

0pNSa7. The efficiency of ultrasonic systems in the case of piezoceramic transducers supplied with rectangular signal. Odobescu G. Liviu (Inst. of Solid Mech., St. Constantin Mille, nr. 15, sector 1, 70701 Bucharest, Romania, odobescu@mecsol.ro)

The paper presents a method for the rise efficiency of ultrasonic systems—in the range 20–100 kHz—supplied by rectangular generators. The rise efficiency is important because of high power, which is working, and for optimal and reliable use of ultrasound systems. The ultrasonic systems constituted from ultrasonic piezoceramic transducers and acoustics chain (transmission, adaptation, amplifying of ultrasound vibration) are supplied from electronic generators, which works in commutation. The commutation-working regime assures a very high efficiency for an electronic generator, but generates a very large spectrum of frequency. The theoretical and experimental analysis of generated spectrum accomplished in this paper leads to very useful conclusions in order to choose an adequate matching dipole between generator and charge. Also, this analysis permits the determination of efficiency for whole ultrasonic system and it found some possibilities for it raising. This study is accomplished by the calculation programs: the efficiency as a function of harmonics, with and without matching dipole; the efficiency as a function of the acoustic charge, with matching dipole, for fundamental frequency; the input impedance of a piezoceramic transducer as a function of acoustic charge; and the total impedance of a piezoceramic transducer with matching dipole as a function of n harmonics.

SUNDAY AFTERNOON, 3 DECEMBER 2000 PACIFIC SALON F, 1:20 TO 2:40 P.M.

Session 0pNSb

NOISE-CON and Noise: Active Noise and Vibration Control

James K. Thompson, Chair

Link Engineering, 1213 Waterways Drive, Ann Arbor, Michigan 48108

Contributed Papers

1:20

0pNSb1. Active increase of sound transmission loss of partitions based on utilizing EMFi panel actuators. Hannu Nykanen, Marko Antila (VTT Automation, Safety Eng., P.O. Box 1307, FIN-33101 Tampere, Finland, hannu.nykanen@vtt.fi), Seppo Uosukainen (VTT Bldg. Technol.), and Jukka Lekkala (VTT Chemical Technol.)

Active noise control (ANC) has been a promising method for noise control and also for other applications such as active increase of absorption and active increase of sound transmission loss of partitions. However, only a few applications have been commercially attractive. At the Technical Research Centre of Finland (VTT) a method to manufacture thin, flat acoustic sensors and actuators (flat microphones and loudspeakers) utilizing ElectroMechanical Film (EMFi) technology has been developed. Flat sensors and actuators can be implemented in ANC applications, where large surfaces should be covered with sensors or anti-sound producing elements. EMFi sensors and actuators are typically 4- to 10-mm-thick elements. They can also be collocated, forming an advanced sensor–actuator combination. Results of research focusing on the use of EMFi technology to increase the sound transmission loss of partitions by active means are presented. By active means the sound transmission loss of the partitions against broadband noise could be increased up to 10 dB in the frequency range 100–350 Hz. Results are new and of great importance, because new ANC technology is being developed. The commercial production of EMFi panel actuators will start in the middle of the year 2000.

1:40

0pNSb2. Robust active control of a vibrating plate. Stephen Prajna[a] (Control and Dynamical Systems, California Inst. of Technol., Pasadena, CA 91125, prajna@ieee.org), Oliver Kaiser (ETH—Swiss Federal Inst. of Technol., CH-8092 Zurich, Switzerland), Stanislaw Pietrzko (EMPA—Swiss Federal Lab. for Material Testing and Res., CH-8600 Duebendorf, Switzerland), and Manfred Morari (ETH—Swiss Federal Inst. of Technol., Automatic Control Lab., CH-8092 Zurich, Switzerland)

This paper addresses several aspects of designing robust controllers for active control of vibration in a plate with piezo actuators and acceleration sensors. The discussion ranges over identification of the plant, design and implementation of controllers, and evaluation of their robustness and performance. All the concepts are tested on an experimental setup. Robustness issues are prominent, since in practice the plant differs from models obtained through identification, and moreover changes in ambient conditions induce variation over time in the plant's characteristics such as modal frequencies and damping ratios. Consequently, a controller designed for a nominal plant model may yield a poor performance or even result in an unstable system if robustness issues are not considered properly. Several H_∞ controllers are compared here. Besides using the usual multiplicative and additive uncertainty models, controllers are also designed using the eigenvalue perturbation approach. It is found that controllers designed using this approach are more robust than those obtained using conventional methods. Furthermore, the performance does not degrade severely (which usually occurs due to the robustness performance

tradeoff). Hence, it is concluded that the approach is suitable for active vibration control of a plate or similar plants. [a] Previously with Fac. Mathematical Sciences, Univ. Twente, The Netherlands.

2:00

0pNSb3. Active control of rotating shafts using modal separation. D. C. Swanson and B. Kim (ARL Penn State, Rm. 9 ASB, P.O. Box 30, State College, PA 16804, Dcs5@psu.edu)

Modal filtering and separation is used to process error signals of the orbital displacement of a high-speed rotating shaft. The error signals are sampled as complex numbers at two shaft positions where the horizontal displacement is the real part and a vertical displacement is the imaginary part. Two complex error signals are then obtained from two separate positions along the rotating shaft. In the frequency domain, the modes of the rotating shaft can be separated as principal components of an ellipse, with spatial phase defined relative to a tachometer signal from the shaft. The active control system actuators (vertical and horizontal directions at two shaft positions) can then be independently driven to minimize each mode. [Work sponsored by ONR.]

2:20

0pNSb4. Robust control of sound in an active acoustic grillage. Kalev Sepp and Juris Vagners (Dept. of Acrosp. and Astron., Univ. of Washington, Box 352400, Seattle, WA 98195)

This study evaluates the results of robust control of an active acoustic grillage (AAG). The AAG is an array of sound sources and microphones forming an acoustic barrier trapping sound but allowing air and light to pass through. Due to strong coupling between the cells of the AAG and resulting acoustic ''feedback,'' a multivariable H_∞-synthesis is used for the controller design to yield a stable and robust closed loop system. The multi-input–multi-output (MIMO) design is compared with a linear-quadratic-Gaussian (LQG) design and a classical SISO approach that neglects the cross-coupling between the cells. System modeling, H_∞-synthesis including the weight selection, simulation and implementation issues are discussed. The H_∞-controller is tested on a four-cell grillage installed on the rectangular duct with cross section 50×50 cm. The robust controller provides 10–12-dB attenuation in the frequency range of 100–300 Hz that is superior to performance achieved with LQG or SISO methods. The controller implementation includes additional features like calibration and real-time monitoring. The work demonstrates benefits of robust MIMO design in noise control systems with multiple sensors and actuators where cross coupling is significant.

SUNDAY AFTERNOON, 3 DECEMBER 2000 PACIFIC SALON F, 2:40 TO 4:00 P.M.

Session 0pNSc

NOISE-CON and Noise: Aircraft Noise

Alan H. Marsh, Chair

DyTec Engineering, Inc., 16072 Santa Barbara Lane, Huntington Beach, California 92649-1213

Contributed Papers

2:40

0pNSc1. Modeling the normal modes and acoustics of a jet engine. Laralee G. Ireland and Scott D. Sommerfeldt (Dept. of Phys. and Astron., Brigham Young Univ., N283 ESC, Provo, UT 84062, ireland@physics.byu.edu)

Fan noise is a major contributor to the noise problem of an airplane jet engine. An understanding of the acoustics of the jet engine is necessary in order to successfully control the fan noise of the engine. A numerical boundary element method model has been developed using Sysnoise to simulate active noise control of fan noise in the engine duct. Control of multiple radial modes is targeted through the use of a modal controller using Newton's algorithm. Results from the model are compared with measurements made previously using the NASA Glenn Research Center test rig facility. Currently work is progressing on the numerical implementation of control on a more realistic geometry.

3:00

0pNSc2. Survey of human response to control of low-frequency tiltrotor interior noise. Scott M. Hirsch (Boeing Aircraft and Missile Systems, P.O. Box 16858, MS P32-74, Philadelphia, PA 19142, scott.m.hirsch@boeing.com)

A preliminary subjective test is conducted in which subjects rank their preference of sounds simulating the interior noise field of a large tiltrotor in forward flight. This noise field is characterized by large-amplitude tonal penetration at the blade passage frequency (<20 Hz) and its first seven overtones. Proposed noise control efforts are simulated by reducing the amplitudes of some or all of these tones. Human response to variations in three noise control parameters is studied: the number of harmonics which are reduced in amplitude, the frequency(ies) of the harmonics which are reduced in amplitude, and the amplitude by which the harmonics are reduced. The results of this test demonstrate that an effective noise control strategy must reduce multiple harmonics, focusing on those which are heard most easily (either the loudest or those with the highest frequency). It is not necessary to control the lowest harmonics, which impose the greatest weight and bulk penalties on a noise control system. It is demonstrated that a harmonic amplitude reduction of only 6 dB may be sufficient to significantly improve passenger response. The results of this test can be used to guide current low-frequency noise control research and development efforts.

3:20

0pNSc3. Correlations in fuselage response and sound field in aircraft cabin during flight. Gopal P. Mathur, Boi N. Tran, and Myles A. Simpson (The Boeing Co., 2401 E. Wardlow Rd., MC: C078-0420, Long Beach, CA 90807)

The in-flight acoustic field in the aircraft cabin is of fundamental interest to the noise control engineer as it defines the sound quality of the cabin environment which the passenger experiences. Design of efficient noise control treatments is usually based on the spectral characteristics of the interior acoustic field to provide a comfortable environment with a minimum of added weight. In-flight noise diagnostics methods, such as sound intensity mapping, are sensitive to the reverberant nature of the cabin sound field. The performance of active noise control systems is crucially dependent on the statistical characteristics of the enclosed field. However, statistical modeling to characterize the aircraft cabin noise field, e.g., correlation modeling, has received little attention. Apparently, acoustic correlation models used in architectural acoustics have been presumed

to represent the in-flight cabin noise field. The statistical description of the aircraft cabin acoustic field can help in understanding and describing the intricate nature of the noise field in greater detail. The main objective of this paper is to discuss correlation models between the exterior pressure field, the vibration response of fuselage panels, and the interior noise field. Test data from ground and flight tests conducted on an MD-90 aircraft are used to examine these correlations.

3:40

0pNSc4. Evaluation of efficient passive damping treatments for aircraft cabin noise control. Gopal P. Mathur, Joshua T. Lee, and Myles A. Simpson (The Boeing Co., 2401 E. Wardlow Rd., MC: C078-0420, Long Beach, CA 90807-5309)

Visco-elastic damping has been used as one type of passive noise control treatment to reduce noise transmission through fuselage skin panels. In most conventional applications, uniform coverage of the damping layer is applied to skin panels. Recently, stand-off damping has also been used for aircraft noise reduction. Effects of various damping treatments have been well documented in the published literature. In this paper, some weight-efficient damping treatments, such as edge and wedge damping, are considered. The main objective of this research work is to evaluate passive noise control treatments for a stiffened sidewall section using analytical models. The goal is to achieve an improvement of 3–5 dB in the transmission loss of the sidewall over a conventional passive treatment, over the frequency range from 400 to 2000 Hz. The effect of various damping treatments on the vibration response of a stiffened panel to a point force input, acoustic pressure field and turbulent boundary layer excitation were investigated using a NASTRAN spectral model. The finite-element model of the stiffened panel included visco-elastic damping with and without constraining layer. The material properties of a commercially available constrained layer foam damping were used.

SUNDAY AFTERNOON, 3 DECEMBER 2000 PACIFIC SALONS C AND D, 1:20 TO 3:40 P.M.

Session 0pNSd

NOISE-CON and Noise: Community Noise

Gordon L. Bricken, Cochair
Gordon Bricken & Associates, Inc., 1621 East 17th Street, Suite K, Santa Ana, California 92701-8518

Rob Green, Cochair
URS Greener Woodward-Clyde, 2020 East First Street, Suite 400, Santa Ana, California 92705

Contributed Papers

1:20

0pNSd1. The community impact of rogue transportation sources. Ithan Zimmer, Charles Wilson, and Robert Dresnack (New Jersey Inst. of Technol., University Heights, Newark, NJ 07102-1982, ithan.zimmer@njit.edu)

This study investigates the contributions of atypical transportation noise sources to community noise and relates these contributions to policymaking and resource allocation. Small percentages of aircraft and surface transportation vehicles make a significant contribution to noise levels as measured by day-night sound level and other descriptors. Rogue sources are loosely defined in this study as anomalous sources that have a large impact on community noise. Included are sources that do not comply with regulatory standards due to intentional modifications or due to failure of noise suppression components. Also included are sources that produce noise levels well above reasonably attainable levels despite attainment of regulatory standards. Rogue source identification and assessment can be a component in policymaking and resource allocation. Cost savings can be realized by reducing noise at the source vice by the path. For instance, surface transportation noise standards based on best-available technology and vigorous enforcement can be more cost-effective than erection of noise barriers. In this case, the cost is borne by the offender rather than by the community at-large. The authors analyze and quantify the noise contribution of rogue sources and suggest appropriate noise mitigation.

1:40

0pNSd2. Pump station noise and vibration in a residential development. David A. Towers (Harris Miller Miller & Hanson, Inc., 15 New England Executive Park, Burlington, MA 01803, dtowers@hmmh.com)

This paper presents a case study of noise and vibration from a water-supply pump station containing four 3600-rpm pump units located within a large housing development. Residents near the pump station complained about the disturbing noise and vibration that they feared could damage their homes. Subsequent measurements demonstrated that vibrations at the nearest homes posed no danger in terms of even minor cosmetic damage, and were in fact below the average threshold of human perception. However, the measurements clearly indicated floor vibrations at a frequency of 60 Hz, corresponding to the pump rotation speed, that varied periodically with time at a rate that depended on the number of pumps operating simultaneously. Furthermore, the floor vibrations were found to correlate directly with ''ground-borne noise'' at a frequency of 60 Hz. At the closest home, this tonal and time-varying noise was loud enough to be disturbing to the residents in some cases. It was concluded that this condition resulted from the rigid mounting of the pumps combined with efficient vibration propagation through rocky ground to the slab foundation of the home. Thus, vibration isolation of the pump units was recommended to eliminate the ground-borne noise problem.

2:00

0pNSd3. Strategies to evaluate aircraft routing plans at Newark International Airport. Robert Dresnack, Ithan Zimmer, and Charles Wilson (New Jersey Inst. of Technol., University Heights, Newark, NJ 07102-1982, robert.dresnack@njit.edu)

In early 1998, the authors were retained by the New Jersey Department of the Treasury to identify a clear set of principles, tools, resources, and methodologies that could be used to evaluate effectively all aspects of an airport routing plan. The project was initiated at the urging of citizens who had argued for ten years that ocean routing of aircraft at Newark International Airport would substantially reduce associated noise impacts in the subject area. The Port Authority of New York and New Jersey, the Federal Aviation Administration, and Continental Airlines, the major carrier at Newark, were steadfast in their belief that developing an ocean routing plan would create additional flight delays at the airport. Newark Airport has already experienced the greatest delays of any airport in the United

States 9 out of the last 12 years. The plan, they further argued, would increase noise in communities in proximity to the ocean. The authors met with the various stakeholders, reviewed their positions and their consultant's reports, developed independent research, and published a final report in January 1999. The report presented 14 findings and associated recommendations that are the focus of this paper.

2:20

0pNSd4. Assessing the community impact of aircraft noise. Charles E. Wilson, Robert Dresnack, and Ithan Zimmer (New Jersey Inst. of Technol., University Heights, Newark, NJ 07102-1982, chas3325@aol.com)

Environmental impact statements and reports are often required for proposed airport improvements and aircraft routing changes. Direct monetary costs and benefits are usually addressed. But the social cost of noise impact is seldom expressed in monetary form. Cost/benefit ratios are of questionable value when noise impact is not evaluated. As a result, policy decisions and resource allocation may be used on incomplete data. Even when noise data are provided in environmental impact statements, laymen may find the data difficult to interpret. A method is suggested for assessing noise impact in monetary form. A procedure for interpreting day-night sound level contours in terms of probable speech interference is also proposed. Impacts on real estate values and other social costs of noise were estimated in one area. The results suggested that annual ''losses'' due to aircraft noise are significant in cost/benefit analysis. It was observed that air travel and air freight enjoy substantial direct and indirect subsidies including direct appropriations, airport payments in lieu of tax and rent, and tax-free bonds. Inclusion of social cost estimates in environmental impact statements provides a basis for allocation of direct and indirect subsidies, and allocation of funds for noise mitigation.

2:40

0pNSd5. Low-frequency aircraft noise and its effects on residential land use. Andrew Harris (Andrew S. Harris, Inc., 19 University Ln., Manchester, MA 01944, andy.dreamer@gateway.net)

Low-frequency aircraft noise is most noticeable to the rear and side of aircraft that are taking off and are on or near to the ground. Recently concluded at Minneapolis–St. Paul International Airport (MSP), this study investigated levels and effects of low-frequency noise in residential areas. The study built on work conducted at other airports. It reached five significant findings: (1) the primary effect of low-frequency noise annoyance caused by noise-induced rattling of windows, doors and furnishings; (2) the best predictor of annoyance was the average of maximum levels in the 25–80-Hz bands (called the low-frequency sound level, LFSL); (3) the existing residential sound insulation program at MSP does not measurably change the low-frequency noise reduction (LFNR), but it reduces the annoyance as much as a 5-dB reduction in low-frequency noise level; (4) the maximum practical level of noise in new single-family residences is approximately 25 dB, about 10 dB higher than the LFNR found in the typical untreated residence, and (5) the maximum exterior LFSL compatible with residential use is 87 dB.

3:00

0pNSd6. The noise produced by harbor infrastructures. Salvatore Curcuruto, Alberto de Leo, and Concetta Fabozzi (Natl. Agency for Environ. Protection of Italy (ANPA), Via V. Brancati 48, 00144 Rome, Italy)

The Italian legislation regarding acoustic pollution is slowly going to be completed with the issue of all decrees in application of the framework law of 26th October no. 447/95. One of these decrees refers to the noise pollution produced by harbor infrastructures. There is little national and international information regarding this argument, above all because such infrastructures are highly complex, characterized by many acoustically different variables, represented by the single activities that are present in the infrastructure during the daily work. For this study seven Italian harbors were chosen, six marine and one lacustrine. For the choice of the harbors several criteria were used based on importance, territorial characteristic (e.g., characteristic of the harbor area, distance from inhabited areas), and presence of other activities (commercial, tourist and industrial). The first step of the work was the location of adequate cartography of the different noise sources present in the harbors and of the mainly exposed receivers. In a second time some measurements of noise were executed outside the harbor, in correspondence to the perimeter and at different distances along a determined directrix, and near the receivers mainly exposed. The results are presented in the present document.

3:20

0pNSd7. Thirty years of traffic-noise problems and countermeasures taken in Greater Cairo, Egypt. S. A Ali and A. Tamura (Dept. of Architect., Faculty of Eng., Yokohyama Natl. Univ., 79-5 tokiwadai, Hodogaya-Ku, Yokohama City, 240-8501 Japan, sayed102@hotmail.com)

Several efforts connected with traffic-noise problems in Greater Cairo have been conducted since the 1970s. They concerned both the traffic-noise measurements and countermeasures taken. During 30 years, the level of traffic noise has been changing, it increased seriously in the last 10 years in spite of countermeasures taken, due to a significant increase in the number of cars. Measurements taken for traffic-noise levels indicated that noise levels in the city are too high. So that makes the city in the front of Capitals from aspect of noise level and continuous periods. In our paper, we have presented measurements and countermeasures taken as follows: (1) We present 30 years of traffic-noise problems in Greater Cairo and the measurements taken. (2) We present the policies for traffic-noise problems and countermeasures taken; this contains the efforts of the Egyptian Ministry of Environment, City Hall, traffic administration of the city, recommendations of the Symposium of Noise and Traffic Problems in Greater Cairo 1988, and recommendations of the Report of the Egyptian Academy of Scientific Research and Technology, and the Egyptian National Research Center 1994. This is to evaluate this experiment, and address planning for the future to make Greater Cairo a quiet city. (3) Conclusions and remarks will be discussed.

Session 0pNSe

NOISE-CON and Noise: Plenary Session—Global Players in Acoustic Standardization

Alan H. Marsh, Chair
DyTec Engineering, Inc., 16072 Santa Barbara Lane, Huntington Beach, California 92649

4:20

0pNSe1. Global players in acoustic standardization. Klaus Brinkmann (Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany, klaus.brinkmann@ptb.de)

The World Trade Organization has identified international standardization as a key to realization of worldwide free trade. This identification underlined the importance of the work by nongovernmental international standardization bodies such as IEC and ISO as well as by governmental bodies such as the International Organization of Legal Metrology (OIML). In acoustics, each of these organizations deals with technical harmonization in a certain field: The work of both IEC/TC 29 ''Electroacoustics'' and OIML/TC 13 ''Measuring instruments for acoustics and vibration'' focuses on specification of performance characteristics and testing procedures for electroacoustical devices. Standardization of measuring methods for industrial products either emitting or reducing noise is the primary activity of ISO/TC 43 ''Acoustics.'' This lecture presents an introduction to how international standards in acoustics are developed along with a survey of the scopes and work programs of these three closely cooperating committees. Guidance on the estimation of the uncertainty inherent in standardized acoustical tests and measurements is a major technical issue. The role of national and regional standardization in acoustics will be noted. The European standards bodies, CEN and CENELEC, have clearly defined tasks within the European Economic Area related to acoustics.

Session 1aAO

Acoustical Oceanography: Bioacoustics I

Timothy K. Stanton, Chair

Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Bigelow 201, Woods Hole, Massachusetts 02543-1053

Chair's Introduction—7:50

Invited Paper

8:00

1aAO1. Acoustics in fisheries in the 21st century. George A. Rose (Fisheries and Marine Inst. of Memorial Univ. of Newfoundland, St. John's, NF A1C 5R3, Canada)

The history, current state, and future potential of acoustics in fisheries research and management will be addressed. Fisheries management worldwide is undergoing a fundamental paradigm change in response to population pressures and overharvesting of marine and freshwater stocks. Future fisheries management will be more conservation based and responsive to ecosystem considerations. These trends will demand more comprehensive knowledge of the components of fisheries' ecosystems and their dynamics, and more realistic estimates of the precision of ocean measurements. Acoustics is poised to take a new and leading role in these developments. Acoustics is nondestructive of marine life, can be near-synoptic, and can measure most of the water column. However, acoustic methods are still hampered by problems with detectability, especially near boundaries, uncertainties about target strengths of the myriad of species and life stages whose scattering properties are complex and behaviorally mediated, and taxonomic identification. These problems will be discussed, and a general experimental plan will be laid out to solve them. The 21st century of fisheries belongs to acoustics and allied technologies.

Contributed Papers

9:00

1aAO2. Backscatter measurements of Atlantic herring in the Northwest Atlantic. William L. Michaels, J. Michael Jech (Northeast Fisheries Sci. Ctr., 166 Water St., Woods Hole, MA 02543), and David A. Demer (Southwest Fisheries Sci. Ctr., La Jolla, CA 92038)

Annual fisheries acoustic surveys have recently been implemented in the Gulf of Maine and Georges Bank regions to estimate the abundance and biomass of Atlantic herring (*Clupea harengus*). Research was devoted to the verification of backscatter measurements from Atlantic herring using multifrequency echo-integration, omni-directional sonar, pelagic trawling, and underwater video operations. Results indicated that individual backscatter measurements during the day were significantly higher compared to the night and twilight periods. The diurnal behavior patterns of Atlantic herring contributed to this variability because of their vertical migration to near bottom during day and to surface waters during night. Observations showed that herring shoaled in stationary schools above offshore banks during the night to feed, and tend to swim more actively during twilight and day periods. An increase in the backscatter from herring during the day may be explained by their more horizontal orientation when actively swimming. Fisheries acoustic surveys conducted during day and night must incorporate systematic adjustments to account for the diurnal bias in acoustical estimates.

9:15

1aAO3. Broadband acoustic backscattering from alewife fish: Experiment and analysis. D. Benjamin Reeder, Timothy K. Stanton, Dezhang Chu (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, breeder@whoi.edu), and J. Michael Jech (NOAA/NMFS Northeast Fisheries Sci. Ctr., Woods Hole, MA 02543)

A series of laboratory scattering measurements and associated analysis have been conducted involving alewife fish (Alosa pseudoharengus). A greater-than-octave bandwidth, shaped, linearly swept, frequency modulated signal was used to insonify live, adult alewife that were tethered while being rotated. The acoustic scattering time series were measured in 1-deg increments of orientation angle over all angles in two planes of rotation. Dominant acoustic scattering mechanisms are identified through both spectral and time-domain (pulse compression) analyses. Results demonstrate the dependence of scattering strength upon frequency and angle of orientation in the lateral (horizontal) and dorsal/ventral aspects. The pulse compression processing of the echoes from the animals temporally resolves multiple returns from an individual which are correlated with size and orientation. [Work supported by ONR, NOAA and CICOR (Cooperative Institute for Climate and Ocean Research).]

9:30

1aAO4. Three-dimensional visualization of acoustic backscattering and models by alewife. J. Michael Jech (NOAA/NMFS Northeast Fisheries Sci. Ctr., 166 Water St., Woods Hole, MA 02543), D. Benjamin Reeder, Timothy K. Stanton, and Dezhang Chu (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

The advent of nontraditional underwater acoustic instrumentation (e.g., multibeam and sector-scanning sonar) and incorporating fish behavior in quantitative fisheries assessment require that backscatter be measured and modeled at a variety of aspect angles. Acoustic backscattering measurements were obtained in a large laboratory tank using live alewife (*Alosa pseudoharengus*). An individual alewife for each series of measurements was tethered and rotated in two planes of orientation (dorsal/ventral and lateral). The alewife were insonified with a broadband (40–100 kHz) chirp signal and bistatic scattering geometry was used. Backscattering amplitudes for all angles of orientation (3-D scattering ambit) were modeled using a Kirchhoff ray-mode model [C. S. Clay and J. K. Horne, J. Acoust. Soc. Am. **96**, 1661–1668 (1994)] and digital images of the fish body and swimbladder morphometry. Visualization of the scattering ambit provides a quantitative examination of the effects of fish orientation on echo amplitude. Comparisons of backscattering amplitudes from the model and measurements along the dorsal/ventral and lateral planes are given. The

utility of broadband measurements for fish backscattering amplitude measurements and the integration of acoustic models in fisheries assessments are discussed. [Work supported by ONR, NOAA/NMFS, and CICOR.]

9:45–10:00 Break

10:00

1aAO5. Incorporating behavior in backscatter model predictions of walleye pollock target strength. John K. Horne (Univ. of Washington and Alaska Fisheries Sci. Ctr., 7600 Sand Point Way NE, Bldg. 4, Seattle, WA 98115, john.horne@noaa.gov) and Michael W. Davis (Fisheries Behavioral Ecology Program AFSC, Hatfield Marine Sci. Ctr., Newport, OR 97365)

Target strength to size relationships for fish and zooplankton are typically based on *in situ* field or *ex situ* tethered measurements. Combining backscatter measurements with species-specific backscatter models improves the translation of acoustic data to computations of type, number, and size of aquatic organisms. Variability in acoustic measurements is largely due to the choice of carrier frequency and organism behaviors. Shoaling, schooling, and orientation are three behaviors that influence the amplitude of echoes received from individuals or aggregations of aquatic organisms. Kirchhoff-ray mode (KRM) backscatter models of individual fish were combined with tilt and roll tank observation data to predict and visualize backscatter from individual and groups of walleye pollock (Theragra chalcogramma). Model predictions were compared to *in situ* target strength measurements at 38 kHz in the Bering Sea. Incorporating behavior in distributions of echo amplitudes should increase accuracy of target strength to length and abundance estimate accuracy. [Work supported by ONR and the Alaska Fisheries Science Center.]

10:15

1aAO6. Capelin TS: Effect of individual fish variability. Yvan Simard (Dept. of Fisheries and Oceans, Maurice Lamontagne Inst., Mont-Joli, QC G5H 3Z4, Canada), John K. Horne (Univ. of Washington, Seattle, WA), Diane Lavoie, and Ian McQuinn (Maurice Lamontagne Inst., Mont-Joli, QC, Canada)

Capelin (*Mallotus villosus*) is an important forage fish in northern latitudes. The effect of the individual biological variability on the target strength (TS) at various acoustic frequencies was investigated from a sample collected in the St. Lawrence estuary. The geometric properties of the fish and its swimbladder were measured from radiographs obtained from 45 anesthetized fish, from 12- to 16-cm total length. The data were input to a backscattering model exploring the effect of fish shape on TS as a function of acoustic frequency, length, and tilt angle. The swimbladder had similar cross sections in both lateral and dorsal views. It represented 5.5% (s.d. 1.1%) of lateral body cross section and 8.2% (s.d. 1.8%) of the dorsal body cross section. The swimbladder cross section was related to the fish total length but the variation around the mean for a given length was $\pm 40\%$. This large variability is equivalent to the change in cross section between tilt angles of 0 and 90 deg for an average fish. The variability in shape parameters is paralleled with changes in the modeled backscatter patterns. The TS versus frequency relation exhibits substantial peaks and troughs, notably in the range of acoustic frequencies commonly used in fisheries acoustics ($\sim$38–200 kHz).

10:30

1aAO7. *In situ* target strength measurements of a deep-water fish, orange roughy. Rudy J. Kloser (CSIRO Marine Res. Hobart, Tas, Australia 7001)

In situ target strength methods have been reported to be the best way to convert echo integration acoustic surveys into biomass. Detailed *in situ* target strength measurements were obtained by both drifting over and towing through aggregations of orange roughy at 700–800-m depths. The aggregations displayed a marked avoidance response at times to the deep-towed body and it was difficult to ascertain when single orange roughy targets were being detected. Using three frequencies, it was possible to include/exclude various targets in the selection process and have more confidence that the observed targets were the species of interest. Calibration of the data was simplified by suspending a sphere under the transducer for all the drift experiments. The extracted target strength histograms for the drift experiments were very broad with many modes, and distribution mean target strengths were sensitive to the density of the schools and also not surprisingly the concentration of other species. Towing experiments were highlighted by a marked avoidance response as the towed body neared the orange roughy schools. The extracted *in situ* target strength histograms showed that the mean target strength for diving orange roughy can decrease by 3.1 dB (50%).

10:45

1aAO8. Some problems and solutions for the measurement of fish target strength: A study case with Atlantic redfish (*Sebastes spp.*). Stéphane Gauthier and George A. Rose (Fisheries Conservation, Marine Inst., Memorial Univ. of Newfoundland, P.O. Box 4920, St. John's, NF A1C 5R3, Canada)

Potential biases in the measurement of redfish target strength (TS) were examined by comparing *ex situ* and *in situ* approaches. *Ex situ* experiments were conducted on individuals after developing a method to maintain live fish in good condition. The main problem with the manipulation of these species was the inflation and distortion of the swimbladder. To minimize this bias, fish were kept in sea cages after acclimation at depth and TS was measured in a camera-monitored apparatus set in proximity to the capture site. A series of *in situ* acoustic-trawl experiments was conducted on several aggregations of redfish in Newfoundland waters. TS were collected using a hull-mounted EK500 split-beam transducer and a deep-tow dual-beam system. The dual-beam transducer was calibrated at different depths to test if change in pressure and/or temperature affected its sensitivity. This system was used to measure aggregations of fish under different transducer depth. The data indicated that biases in *in situ* TS estimates increased with the range of observation, the density of fish, and the presence of multiple targets formed by the clustering of smaller organisms. Depending on the nature of the bias, TS can be over- or underestimated by as much as 6 dB.

11:00

1aAO9. The development of a system for remotely monitoring the three-dimensional movement of acoustically tagged fish. Tracey W. Steig and Samuel V. Johnston (Hydroacoustic Technol., Inc., 715 NE Northlake Way, Seattle, WA 98105, consulting@htisonar.com)

Over the last 3 years, a passive acoustic system was developed to monitor the behavior of acoustically tagged fish. The acoustic tag receiver was designed to simultaneously monitor up to 16 omni-directional hydrophones. The signals received at the hydrophones were synchronized to determine the arrival times for each pulse transmitted by the acoustic tag. Arrival times were used to calculate the three-dimensional position of a tagged fish as it moved through the water. Since precise three-dimensional positions of the hydrophone array was required, methods were developed to accurately determine these positions. Additional algorithms were developed to track the received signals and calculate the three-dimensional position of a tagged fish. The resulting acoustic tag system was deployed at various hydroelectric dams on the Columbia, Snake, and Cowlitz rivers. Field studies monitored downstream migrating salmonids as they approached and passed into the turbine intakes, spillways, and juvenile bypass channels. The size of the juvenile salmonids ranged from 160 to 240

mm. The tags were approximately 7 mm in diameter, 23 mm in length, and operated at a transmitting frequency of 300 kHz. Fish movement patterns were tracked over time, with submeter resolution in most cases.

11:15

1aAO10. A method for estimating the accuracy of the three-dimensional position of acoustically tagged fish. John E. Ehrenberg and Tracey W. Steig (Hydroacoustic Technology, Inc., 715 NE Northlake Way, Seattle, WA 98105, consulting@htisonar.com)

Acoustic tag systems are beginning to be used to study the behavior of fish in a region of interest. In particular, tag systems are being successfully used to study the behavior of downstream migrating salmon smolts (Oncorhynchus spp.) as they approach hydroelectric dams. While field studies have demonstrated the potential for acoustic tag systems, very little has been done to quantify the performance of these systems. This paper develops a method for predicting the accuracy of the position estimates provided by acoustic tag systems. General expressions are developed that can be applied to any particular deployment of a tag system. In addition, some specific examples are used to provide general guidelines that should be followed to achieve good performance when deploying an acoustic tag system.

11:30

1aAO11. An active fish tracking split-beam sonar to study salmon smolt behavior. John B. Hedgepeth (Tennera Environ., LLC, 3427-A Miguelito Court, San Luis Obispo, CA 93401) and Gary E. Johnson (BioAnalysts, Inc., Battle Ground, WA 98604)

An active fish tracking split-beam sonar (AFTS) is being used to study salmon smolt behavior at the Dalles Dam on the Columbia River. AFTS is based on the principle of tracking radar. Once a smolt is detected, two high-speed stepper motors align the axis of a digital split-beam transducer on the target. As the fish moves, deviation of the target from the beam axis is calculated and used to re-aim the transducer, thereby tracking the target. For each ping the target is tracked, data on fish position and target strength are recorded to disk. Fish position resolution is ± 0.5 cm within a sample volume approximately 14 m of range from the transducer. Individual fish tracks are visualized in three-dimensional plotting software. AFTS tracks are coupled with hydraulic data from a computational fluid dynamics model to determine statistical relationships between fish movements and flow conditions. Sequential fish positions are also analyzed using a Markov process. Fish states observed with AFTS include holding, swimming actively downstream, and passive drift. The information from AFTS will aid design of bypass systems to protect endangered and threatened salmon populations in the Pacific Northwest. [This research is funded by the Corps of Engineers.]

11:45–12:15

Panel Discussion

MONDAY MORNING, 4 DECEMBER 2000 PACIFIC SALON A, 7:55 A.M. TO 12:15 P.M.

Session 1aBB

Biomedical Ultrasound/Bioresponse to Vibration: Topical Meeting: Physics of Echo-Contrast Agents I

E. Carr Everbach, Chair
Department of Engineering, Swarthmore College, Swarthmore, Pennsylvania 19081-1397

Chair's Introduction—7:55

Invited Papers

8:00

1aBB1. Modeling of ultrasound contrast agents. John Allen (Dept. of Biomed. Eng., One Shields Ave., Univ. of California, Davis, Davis, CA 95616, jsallen@ucdavis.edu)

Ultrasound contrast agents differ from free-gas bubbles such that a polymer, lipid, or fluid shell encapsulates them. Many modeling efforts have focused on the role of the shell on the overall dynamics. Early efforts used semiempirical approaches to incorporate the shell into a generalized Rayleigh–Plesset equation. Church [J. Acoust. Soc. Am. **97**(3) (1995)] developed a more rigorous approach modeling the albumin shell of Albunex as a linear elastic material. This work has provided the foundation for many subsequent theoretical and experimental efforts. Ye [J. Acoust. Soc. Am. **100**(3) (1996)] examined modal scattering solutions for an Albunex agent. Recent efforts to design a high-frequency contrast agent based on dipole scattering are discussed [Allen, Kruse, and Ferrara, IEEE-Trans. Ultrason. Ferroelectr. Freq. Control, in press]. The modeling of different types of shell materials remains an outstanding issue. The potential role of both viscous and viscoelastic shell materials is outlined. Furthermore, the limitations of linear elastic and viscoelastic material models are highlighted. The modeling of possible interaction effects among groups of contrast agents has received less attention. Also, outstanding issues exist with respect to a rigorous mathematical modeling of contrast agent destruction, drug delivery schemes, and potential bioeffects.

8:15–9:15

Panel Discussion on Modeling of Echo-Contrast Agents

9:15–9:30 Break

9:30

1aBB2. *In vitro* experimental evaluation of contrast agents. Katherine Ferrara (Div. of Biomed. Eng., One Shields Ave., Univ. of California, Davis, Davis, CA 95616)

In this presentation, techniques for *in vitro* characterization of contrast agents will be reviewed. Contrast agents contain air or a lower diffusivity gas, with a shell that can be composed of albumin, polymer, a lipid monolayer, a surfactant or oil. The range of possible shell properties creates a wide range of physical properties, and resulting variations in behavior upon insonation. In order to characterize the agents, several methods have been employed, including acoustical and optical interrogation of single bubbles and ensembles. High-speed optical techniques can be employed to measure the radius-time curves, to determine the mechanisms of destruction, and to detect surface waves. The maximum radius and wall velocity, and mechanisms of destruction, are significantly affected by both the gas and shell properties. Observed mechanisms of destruction include static diffusion, acoustically driven diffusion and fragmentation. These destruction mechanisms are characterized by unique time scales which have been determined through both optical and acoustical methods. The results of acoustical techniques that have been applied to determine the linear resonance frequency and to quantify the scattering and attenuation cross sections will also be presented.

9:45–10:45

Panel Discussion on *In-vitro* Experimental Evaluation of Echo-Contrast Agents

10:45–11:00 Break

11:00

1aBB3. Imaging techniques for ultrasound contrast agents. Michalakis A. Averkiou (ATL Ultrasound, Philips Medical Systems, P.O. Box 3003, Bothwell, WA 98041-3003)

Contrast agents are used today to enhance and aid organ (myocardium, liver, brain, etc.) perfusion studies. From the early days of contrast imaging it was understood that microbubble specific imaging modalities would have to be invented to utilize the contrast effect. The two main characteristics of contrast microbubbles that have defined today's ultrasound modalities are (1) microbubbles are nonlinear scatterers that generate harmonic signals and (2) they exhibit a transient response, i.e., they are destroyed by ultrasound. Harmonic Doppler and imaging techniques were introduced at first to help with the detection of microbubbles in small vessels and capillaries. The harmonic response of bubbles and their Doppler signals increase with increasing acoustic amplitudes (mechanical index, MI), which results in fast bubble destruction. Triggered (intermittent) imaging, insonifying every n cardiac cycles (where $n = 1,2,3, \ldots$), was established to minimize destruction. Triggered imaging is a cumbersome technique. Real-time imaging techniques are now available. They are made possible by low-amplitude (MI) scanning and by more sensitive low MI techniques like pulse inversion and amplitude modulation. Both triggered and real-time techniques will be discussed in this presentation and clinical examples will be presented.

11:15–12:15

Panel Discussion on Imaging Techniques Using Echo-Contrast Agents

MONDAY MORNING, 4 DECEMBER 2000 PACIFIC SALONS C AND D, 9:00 TO 10:00 A.M.

Session 1aNSa

NOISE-CON and Noise: Plenary Session—Noise Control Engineering for the Airport Railway in Hong Kong

Gregory C. Tocci, Chair

Cavanaugh Tocci Associates, Inc., 327F Boston Post Road, Sudbury, Massachusetts 01776

9:00

1aNSa1. Noise control engineering for the airport railway in Hong Kong—Setting the standards. Glenn Frommer (Mass Transit Railway Corp., MTR Tower, Kowloon Bay, Kowloon, Hong Kong, SAR, PROC)

The Lantau and Airport Railway (LAR) is a 34-km electrical railway that connects Hong Kong Island to the new International Airport. Noise control engineering was implemented throughout the design of the project to provide a safe, comfortable, and maintainable environment both for the passengers and railway neighbors in the most cost-effective manner as possible, and minimize land take and land sterilization nearby the railway. To do so required a unique balance of noise and vibration specification and performance in the following areas: (i) a wheel/rail interface design integrating synergistically, the track and rolling stock design, (ii) the use of a noise assurance plan to manage the noise design of the rolling stock, (iii) the use of aggressive specifications for noise

emissions, reverberation times, and building services within stations, (iv) the use of RASTI prediction to ensure proper design and delivery of the public address system, (v) the use of floating track slabs and other trackforms to minimize the transfer of structure-borne noise from the railway platforms to the overlying developments, and (vi) specialist training in acoustics and vibrations. This paper will summarize the key results of the LAR project.

Session 1aNSb

NOISE-CON, Noise and Architectural Acoustics: Heating, Ventilating and Air Conditioning Noise

Gary W. Siebein, Chair
Department of Architecture, University of Florida, P.O. Box 115702, Gainesville, Florida 32611

Contributed Papers

10:20

1aNSb1. Air handler sound power prediction method based on ARI Standard 260. Stephen J. Lind (The Trane Co., 3600 Pammel Creek Rd., La Crosse, WI 54601, slind@trane.com)

A method of predicting air handler sound power based on ratings for a product line is described. The method provides octave band sound power levels based on ratings obtained using Air-Conditioning and Refrigeration Institute (ARI) Standard 260 Sound Rating Of Ducted Air Moving And Conditioning Equipment. Detailed sound power information for HVAC equipment is not always available, but it is important in accurately predicting noise levels in acoustically sensitive spaces. To address this need, a rating program was undertaken using ARI 260. This standard is a reverberant room technique for sound rating ducted air conditioning equipment using a reference sound source substitution method. Since sound travels from the source to receiver along numerous paths, this standard differentiates between sound power emanating from several common paths called components. Components for this project included ducted discharge, free inlet plus casing, ducted inlet and casing. The standard provides guidance on adequate number of fan sizes, appurtenances and operating characteristics. The intent of the project was to provide a model to predict sound power by unit size, component, operating condition and unit configuration. Good agreement was found between predicted levels and measured data.

10:35

1aNSb2. Condenser fan noise reduction solutions. Michael Bobeczko (Sound International Technology, 19721 Torres Way, Portolla Hills, CA 92679)

There has been a surge in the building of multimillion dollar homes along the California Coast. Often, large commercial style air conditioners are installed with multiple axial flow propeller condenser fans. They can generate low-frequency tones that annoy highly sensitive neighbors. By replacing 3- to 4-bladed stamped metal fans with a 6-, 9-, and 12-bladed highly efficient airfoil fan with pitch compensation, by reducing motor speeds from 1140 to 870 rpm, and by installing a variable speed controller, sound levels were reduced 14 dB @ 77 Hz as well as reducing energy consumption. This noise reduction solutions have proven to be far more effective than traditional barriers.

10:50

1aNSb3. Sound attenuation of dual-wall, fiberglass-lined, circular ducts with and without airflow. Douglas Reynolds (Ctr. for Mech. and Environ. Systems Technol., Univ. of Nevada, Las Vegas, 4505 Maryland Pkwy., Las Vegas, NV 89154)

A major test program for measuring the sound attenuation characteristics of dual-wall, fiberglass-lined, circular ducts from six different duct manufacturers was completed. Octave band dynamic insertion loss values for octave band center frequencies from 63 to 8000 Hz were measured. The inside diameter of the ducts that were tested ranged from 12 to 48 in. Sound tests were conducted with and without airflow. Six airflow velocities from 1000 to 6000 fpm were examined. Regression analyses were conducted on the measured insertion loss values. Regression equations for the dynamic insertion loss values were developed that can be used to estimate the dynamic insertion loss of dual-wall, fiberglass-lined, circular ducts for zero flow and for duct airflow velocity from 1000 to 6000 fpm. The equations are valid for duct diameters from 12 to 72 in.

11:05

1aNSb4. A procedure for measuring background sound levels in rooms. Jerry G. Lilly (JGL Acoust., Inc., 5266 NW Village Park Dr., Issaquah, WA 98027) and Mark E. Schaffer (Schaffer Acoust., Inc., Pacific Palisades, CA 90272)

Background noise criteria for indoor spaces have been in existence for decades without an associated standard measurement procedure to verify compliance. The lack of a standard measurement procedure can lead to a significant difference in reported sound levels in a room, even from experienced professionals. The proposed procedure is intended for measuring background sound levels from continuous noise sources, and is being recommended to ASHRAE Technical Committee 2.6 for inclusion in the 2003 Applications Handbook. The purpose of the procedure is to standardize the instrumentation and method of determining the background sound levels. The procedure can be used by anyone charged with the task of comparing a room's background noise environment with established noise criteria.

11:20

1aNSb5. On HVAC duct acoustical end reflection. Emanuel Mouratidis (Vibroacoustics/BVA Systems Ltd., 727 Tapscott Rd., Scarborough, ON M1X 1A2, Canada, emouratidis@vibroacoustics.com) and Richard J. Peppin (Scantek, Inc., Silver Spring, MD)

Duct end reflection (ER) is the apparent loss of sound power resulting from an abrupt change in a cross-sectional area of the duct. In most references, the magnitude of ER is only a function of frequency and duct size, because of assumed plane-wave propagation. If, between the source and duct termination, there are short duct runs, 90-deg duct bends, offset ducts, T-sections, various end diffusers, high aspect ratio cross sections, the presence of airflow, etc., the relative effects on ER are unknown. The American Society of Heating, Refrigeration, and Air-Conditioning Engineer's Applications Handbook presents tables that provide the power loss due to ER. However, the assumptions used and the limitations of the table are lacking. This paper describes results of laboratory-measured changes in reverberation room sound-pressure level (SPL) created by common system variables that are associated with ER: duct length between noise source and reverb room termination, and several duct termination condi-

tions were investigated using small diameter duct. In a controlled laboratory environment, these system conditions were found to create an apparent random variability (up to 10 dB) in the resulting SPL, for both a high impedance fan reference sound source and a low impedance loudspeaker source.

11:35

1aNSb6. Estimating community noise levels from outdoor condensing units. Gary W. Siebein and Hyeong-seok Kim (Dept. of Architecture, Univ. of Florida, P.O. Box 115702, Gainesville, FL 32611-5702)

Noise levels propagated from cooling towers and condensing units located in partial or fully enclosed equipment courtyards onto adjoining residential properties were studied in this paper. A combination of field measurements at existing installations, computer model studies of the installations and physical acoustical model studies were conducted to determine the acoustical effects of reflections from enclosing walls of nearby buildings and partial equipment yard enclosures on sound propagation at distances away from the units. Predictions using distance attenuation and the effects of partially enclosing walls (Harris, 1991; Miller, 1981) were used to estimate sound levels at various distances from several existing outdoor condensing units. Acoustical measurements were made at the actual installations within the chiller enclosure and at various distances in the surrounding communities. Each of the actual situations was simulated in model form at a scale of 1:10. The process of calibrating the model to the existing situations will be presented. The scale model data were within 2 dB of the actual insertion losses for full size octave bands from 63 to 2000 Hz for condensing units in the open and in various enclosures.

11:50

1aNSb7. Tips for specifying active duct silencers. Steve Wise (Wise Assoc., 1409 E. Skyline Dr., Madison, WI 53705, stevewise@att.net), Christian Carme (TechnoFirst, 48, Aubagne 13676, France), and Dwight Shackleford (American Aldes, Sarasota, FL 34234)

Low-frequency noise in buildings is receiving greater focus in new noise standards. The extraordinary performance of active duct silencers on these low frequencies provides a powerful new element for fan noise control. Acoustical consultants need to know the most appropriate applications for active duct silencers. The authors have been personally involved in the design and start-up of over 1000 active duct silencers installed in the USA and Europe to silence fans on a variety of HVAC applications. This site experience from office buildings, schools, hospitals, semiconductor manufacturing facilities, auditoria, sound test chambers, and cruise ships—both retrofit and new construction—is summarized to guide acoustical consultants in estimating the performance of active duct silencers. Fan and duct silencer system layouts are reviewed, along with performance case histories.

MONDAY MORNING, 4 DECEMBER 2000 PACIFIC SALONS C AND D, 10:20 TO 11:25 A.M.

Session 1aNSc

NOISE-CON and Noise: Classroom Acoustics

David Lubman, Chair
David Lubman and Associates, 14301 Middletown Lane, Westminster, California 92683-4514

Invited Paper

10:20

1aNSc1. Progress report on I-INCE technical initiative for schoolroom acoustics. David Lubman (David Lubman & Assoc., 14301 Middletown Ln., Westminster, CA 92683-4514), Louis C. Sutherland (Rancho Palos Verdes, CA 90275-3908), and Zerhan Karibiber (Yldz Tech. Univ., Istanbul, Turkey)

Good acoustics is an indispensable requirement for verbal learning, and therefore vital to all knowledge-based societies. The World Health Organization has identified the basic acoustical requirements for verbal learning spaces (Guidelines for Community Noise, April 1999). Some nations are well advanced in meeting these requirements while others may lack local guidelines or other resources. The worldwide acoustical design and noise control communities can help nations to achieve good schoolroom acoustics. A technical initiative for that purpose titled ''Noise and Reverberation Control for Schoolrooms'' was approved by the International-INCE General Assembly at its meeting in December 1999. The initiative is intended to be an internationally coordinated program to assist participating nations with engineering issues associated with achieving satisfactory acoustics in learning spaces. Representatives of various I-INCE member societies appointed to this Technical Study Group will be asked to submit a draft report for review by the I-INCE General Assembly within two years of their first meeting. The report is expected to provide practical guidance to school planners, designers, and code officials. The Group will first convene at the Internoise meeting at Nice, France in August 2000. This presentation highlights the results of that meeting.

Contributed Papers

10:40

1aNSc2. Determination of insertion loss for classrooms at a high school. Michael Greene (URS Greiner Woodward Clyde, 2020 E. First St., Ste. 400, Santa Ana, CA 92705)

The construction of a new freeway adjacent to an existing high school in eastern San Diego County, California, prompted the need for a rigorous analysis of the noise effects on the school. The insertion loss of structures (with windows and doors open and closed) at a high school was measured. The test setup consisted of two commercial-grade loudspeakers mounted atop a manually operated lift, associated amplifiers, pink-noise generator, a real-time noise analyzer, and sound-level meters. Noise levels were measured at equivalent distances outside and inside the rooms of interest, to derive the structure's insertion loss. This was done at incident angles of 30, 45, 60, and 75 deg to the building facade. The resultant data from these measurements required the use of specially designed spread sheets to effectively analyze the present the results. The results of the measurements

indicated that improvements to the older classrooms near the freeway would be necessary in order to meet the interior noise standard for classroom spaces.

10:55

1aNSc3. Rating and ranking university classrooms by acoustical quality. Murray Hodgson (UBC School of Occupational and Environ. Hygiene, 3rd Fl., 2206 East Mall, Vancouver, BC V6T 1Z3, Canada)

Nonoptimal classroom acoustical conditions affect speech perception and learning by students. The research reported here developed methods for predicting speech intelligibility in occupied classrooms. These were used to predict speech intelligibility at various positions in 280 university classrooms when 70% occupied, for two instructor voice levels. Classrooms were classified and rank ordered by acoustical quality, as determined by the classroom-average speech intelligibility. Classrooms with poor and bad acoustical quality were identified, so that they could be prioritized for acoustical renovation. The conclusion was that 90% of the 280 classrooms have good or very good acoustical quality with a typical instructor. However, 27 (10%) of the classrooms had poor quality, and one had bad quality. Most rooms were very good at the front and good or very good at the back. With a quiet female instructor, classrooms were poor or fair on average. Most were fair at the front, and fair, poor, or bad at the back. New classroom acoustical design and renovations should focus on limiting background noise from the ventilation systems. They should promote high instructor speech levels at the back of classrooms. This involves limiting the amount of sound absorption that is introduced into classrooms to control reverberation.

11:10

1aNSc4. A resource for creating learning environments with desirable listening conditions. Robin Glosemeyer (Russ Berger Design Group, 4006 Beltline Rd., Dallas, TX 75244), Benjamin Seep (Wrightson, Johnson, Haddon & Williams, Dallas, TX 75244), and Robert C. Coffeen (The Univ. of Kansas, Lawrence, KS 66045, rcoffeen@ukans.edu)

A discussion of this recent publication on classroom acoustics by the Technical Committee on Architectural Acoustics of the Acoustical Society of America which is directed towards school planners, school administrators, school faculty, governmental agencies, architects, and others who are engaged in the planning and construction of new classrooms and in the remodeling of existing classrooms. This publication is intended to be used as an aid in the understanding of the elements of desirable listening conditions in school classrooms.

MONDAY MORNING, 4 DECEMBER 2000 PACIFIC SALON E, 10:20 A.M. TO 12:05 P.M.

Session 1aNSd

NOISE-CON, Noise and Engineering Acoustics: Noise Control Materials and Elements

Glenn E. Warnaka, Chair
Future Technologies, 1612 South Allen Street, State College, Pennsylvania 16801

Contributed Papers

10:20

1aNSd1. A finite-element analysis of a smart foam system for passive/active noise control. Cliff L. Chin, Joshua T. Lee, and Gopal P. Mathur (The Boeing Co., 2401 E. Wardlow Rd., MC: C078-0420, Long Beach, CA 90807-5309)

The smart foam noise control system is designed to reduce sound by a combination of the passive absorbing component of the poroelastic foam and the active component of the voltage-driven PVDF (polyvinylidene fluoride) film. The hybrid noise control methodology allows the control of noise over a wide frequency bandwidth. In this paper, the finite-element (FE) method is used to solve the PVDF-foam-structure-acoustic coupled smart foam system in a two-dimensional domain. An FE model of a two-dimensional smart foam actuator with embedded PVDF film will be presented. The effect of foam properties on acoustic impedance and absorption coefficient of the smart foam system is examined using the FE model. The program also predicts the structural vibration of the foam matrix and the acoustic pressure in a cavity when an array of smart foam actuators are driven by control input voltages which are determined by the active control algorithm.

10:35

1aNSd2. Properties and applications of novel acoustical foams. Suresh Subramonian, Mae Drzyzga (Dow Chemical Co., 200 Larkin Ctr., Midland, MI 48674), Sandrine Gilg, Laurent Remy, and Chung Park (Dow Deutschland, Rhinemunster, Germany)

This paper describes the unique properties of a new line of novel acoustical materials that has been developed for sound and vibration management. The materials are macrocellular thermoplastic foams made from polyolefin resin blends. Fire retardant grades have been developed to meet a variety of building and construction fire codes. The foams have several advantages over incumbent materials, including structural integrity, moisture resistance, ease of fabrication/installation and recyclability. This paper also reviews a case study involving the use of the new foams to significantly dampen the sound level of a noisy industrial facility. The sound pressure levels and reverberation decay times were measured in the target area with an integrating, real time $\frac{1}{3}$-oct spectrum analyzer. Architectural acoustical analysis was performed using ray diagrams and Sabine calculations for determining the amount and optimal placement of the absorptive media to reduce the reverberant noise fields and sound reinforcing reflections. The thickness of the acoustic media was determined by matching the peak absorption frequencies of the foam to the peak noise frequencies of the source. After installation of the absorptive media as per specifications, the reduction in noise level was measured to validate the acoustical design.

10:50

1aNSd3. Elimination of structure-borne noise using an insulating material made from recycled rubber. Paul Downey (Dodge-Regupol, Inc., 33 Craighurst Ave., Toronto, ON M4R 1J9, Canada)

A brief overview of the concept of structure-borne noise will be given from both a theoretical and practical perspective. The methods and materials used to insulate several types of structures against vibrations and the resulting structure-borne noise will be explained. A case study approach will be used to demonstrate the results of these installations. The results will be indicated in the appropriate format, including the preceding and resulting sound pressure levels. The background technical data will be presented including the spring characteristics of the materials, the natural frequency, the dynamic stiffness, and the dynamic modulus of elasticity in

dependence of the load. The behavior of the materials under load and at removal of load as well as the effect of layering materials will be examined. The required calculations to permit acoustical engineers to solve structure-borne noise problems will be illustrated.

11:05

1aNSd4. Geometric sound absorbers. Glenn E. Warnaka (Future Technologies, LLC, 1612 S. Allen St., State College, PA 16801-5915, future@vicon.net)

Acoustic absorption is a commonly used technique for noise reduction. Most acoustically absorbent treatments use fibrous or porous materials. In general, such treatments are not rugged and have limitations in many applications. For example, fibrous and porous materials can retain pollutants present in the environment. This leads to unsanitary and unhealthy conditions and to fire and explosion hazards. This paper describes novel geometric acoustic absorbers, based on closely coupled quarter-wave resonators that provide a high level of acoustic absorption over a broad range of frequency. Because the absorbers depend on their geometry for absorption, the absorbers may be made of any material and, hence, the material may be chosen on the basis of cost, environmental resistance, ruggedness, etc. Further, the absorbers can be cleaned when required to eliminate the presence of contaminants. Cleaning can be accomplished by conventional means such as washing, vacuuming, high-pressure air, etc. Two absorbers are described, Configuration ∇ and Flat Absorber. The absorbers may take many forms and may be made compact by folding them or stacking them in various ways without affecting their absorption. In other cases, the absorbers can be made very rugged to withstand heavy use.

11:20

1aNSd5. Broadband structural acoustic silencers. Sripriya Ramamoorthy and Karl Grosh (Dept. of Mech. Eng., Univ. of Michigan, Ann Arbor, MI 48109-2125)

It has long been a goal of the noise control community to achieve broadband passive attenuation of sound propagating in waveguides. By using designs based on models of the mammalian cochlea, theoretical studies show that a structural-acoustic silencer can achieve the desired broadband transmission loss and anechoic termination is possible provided sufficient losses are present. The mammalian cochlea can be idealized as a compact rectangular duct system ($2.5\times0.1\times0.1$ cm) interacting with a strongly orthotropic plate of variable impedance (known as basilar-membrane). A coupled fluid-structure pressure-displacement wave propagates along the basilar-membrane resonating at a frequency-dependent location. In this process, the wave is attenuated with an accompanying reduction in duct pressure (nearly 100 dB over 20–20 000 Hz for humans). Current research intends to imitate this aspect of cochlear behavior for reducing noise level in pipes. Possible applications include use in automotive mufflers and quieting hydraulic systems. Design of fluid/structure interaction between the orthotropic plate and duct is based on numerical modeling using the 2.5-D finite element/analytic technique. Results based on interaction of a varying thickness plate made of typical orthotropic materials like unidirectional graphite-epoxy composite, with fluids like water and oil, are presented. Comparison is made against a constant thickness plate made of the same material.

11:35

1aNSd6. Wideband transmission noise reduction of smart panels featuring piezoelectric shunt circuits and sound absorbing material. Jaehwan Kim, Joong-Kuen Lee, and Jin-Young Choi (Dept of Mech. Eng., Inha Univ., 253 Yonghyun-Dong, Nam-Ku, Incheon 402-751, Korea, jaehwan@inha.ac.kr)

The possibility of a wideband noise reduction of piezoelectric smart panels is experimentally studied. A piezoelectric smart panel is basically a plate structure on which piezoelectric devices with shunt electronics are mounted and sound absorbing material is bonded on the surface of the structure. Sound absorbing material can effectively absorb the sound transmitted at midfrequency, while the use of piezoelectric shunt circuits can reduce the transmission at resonance frequencies of the panel structure. Piezoelectric shunting circuits, composed of register and inductor, are determined based on maximizing the dissipated energy through the circuit. To be able to reduce the sound transmission at low panel resonances, multimode damping with one piezoelectric device is essential and it is achieved by using the measured electrical impedance model. The noise reduction performance of the panels is tested on an acoustic tunnel. The tunnel is made of a guided tunnel that has a square cross section and a loudspeaker is used as a sound source. Panels are mounted in the middle of the tunnel and the transmitted sound pressures across panels are measured. Several panels according to different absorbing material thickness, combination of double panels, and location of piezoelectric devices are tested.

11:50

1aNSd7. Active control of sounds with large dynamic range. Anthony J. Brammer and George J. Pan (Inst. for Microstructural Sci., Natl. Res. Council, Montreal Rd., Ottawa, ON K1A 0R6, Canada)

The performance of an active noise control system employing digital signal processing is influenced by the analog signal amplitudes within the input and output analog-digital (A/D) and D/A converters, and hence is sensitive to the sound pressure being controlled. While compensation is commonly provided within the algorithm for variations in input power (e.g., by normalizing the adaptation step size), the full dynamic range of the A/D and D/A subsystems is usually not realized. An analog gain control system has been developed consisting of linked, reciprocal variable gain amplifiers, so arranged that the changes in signal amplitude at the A/Ds and D/As are smaller than the changes in sound pressure of the acoustic system being controlled. The gain control system operates on the error and secondary source signals so as to maintain the error path impulse response unchanged. In this way the gain changes are transparent to the digital controller. The application of the method to an adaptive feedforward active noise control system for a circumaural hearing protector, or headset, will be described. [Work supported by the Defence and Civil Institute of Environmental Medicine.]

Session 1aSC

Speech Communication: All Things Prosodic (Poster Session)

Amy J. Schafer, Chair
Department of Linguistics, University of California, 3125 Campbell Hall, Los Angeles, California 90095-1543

Contributed Papers

All posters will be on display from 9:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 9:00 a.m. to 10:30 a.m. and contributors of even-numbered papers will be at their posters from 10:30 a.m. to 12:00 noon.

1aSC1. The effect of syllabification and VOT on *F*2 onsets at the CV2 boundary. Golnaz M. Ghavami, Harvey M. Sussman (Dept. of Linguistics, Univ. of Texas, Austin, TX 78712), and Bjorn Lindblom (Univ. of Texas, Austin, TX 78712 and Univ. of Stockholm, Stockholm, Sweden)

The articulatory blending of vowels and consonants has long been described by a coproduction model of coarticulation [S. E. G. Ohman, J. Acoust. Soc. Am. **41**, 310–319 (1967)]. Variation in intergestural timing of the C closure has been shown to produce instability in locus equation coefficients as place descriptors when simulated on the DRM vocal tract model [Chennoukh *et al.*, J. Acoust. Soc. Am. **102**, 2380–2389 (1997)]. This study attempted to achieve real-world variations in CV phasing by imposing a variety of temporal intervals separating vowel-to-vowel sequences. Variation in syllable shapes and degree of aspiration were used to alter the relative timing between V1 and V2. A coproduction account predicts little change in *F*2 onsets for V2 as a function of the temporal interval separating V1 and V2. In contrast, a segment-by-segment (''deactivation'') account would predict systematic changes in V2 locus frequencies, with coarticulation decreasing as the C-interval increases. Three syllable shapes (V#CV, VC#V, VC#CV) were used with five vowel contexts /i,e,ae,u,ɔ/) surrounding the six stop consonants (/bdgptk/). Acoustic measures of both F2 offset frequencies for V1 as well as locus equations derived from the CV2 interface will be presented for three speakers of American English. [Work supported by NIH.]

1aSC2. The effect of stress on *F*2 onsets in VCV utterances. Ramin Sarraf, Harvey M. Sussman (Dept. of Linguist., Univ. of Texas, Austin, TX 78712, ramin.sarraf@mail.utexas.edu), and Bjorn Lindblom (Univ. of Texas, Austin, TX 78712 and Univ. of Stockholm, Stockholm, Sweden)

This study compares a superposition or vowel-to-vowel theoretical framework versus a segment-by-segment account to best account for the coarticulatory dynamics of VCV utterences. The ''trough effect,'' a deactivation of tongue musculature after the initial vowel of a symmetrical VCV sequence (e.g., /ibi/), is not readily explainable by traditional coproduction (vowel-to-vowel) accounts of coarticulation. A Jossian segment-by-segment programming sequence appears to be more compatible with ''deactivation'' phenomena. One way to manipulate extent of ''deactivation'' is to impose differential degrees of syllable stress on either side of the stop consonant. Locus equations can provide an acoustic window on this question as timing differences should show up as slope/intercept variations. Emphatic stress on *V*1 is expected to lessen deactivation relative to weak *V*1 stress, and consequently greater levels of anticipatory coarticulation should occur, resulting in higher locus equation slopes. Four different stress-vowel conditions were used: weak versus emphatic stress counterbalanced over *V*1 and *V*2. There were five vowel contexts for *V*1 and ten vowel contexts for *V*2, surrounding the three voiced stop consonants /b, d, g/. Results will be reported on three speakers of American English. [Work supported by NIH.]

1aSC3. Function word reduction: The role of adjacent stress. Lisa M. Lavoie (MIT, Rm. 36-511, 50 Vassar St., Cambridge, MA 02139, lisa@speech.mit.edu)

Monosyllabic function words in speech are often reduced from their full citation forms. Researchers have shown that reduction is influenced by part of speech, predictability, prosody, and disfluency. This work isolates the role of an additional factor, preceding stress, in function word reduction by analyzing the realization of ''for'' in a highly controlled corpus. Five American English speakers read pairs of disyllabic words with initial or final stress (BOOty, bouTIQUE) in the frame ''Please say [word] for me.'' Two speakers show striking sensitivity to preceding stress, always producing strong–weak alternations, such that strong ''for'' follows a weak syllable and vice versa. The rimes in the strong versions average 90 ms while the weak are just 50 ms. The strong rimes consist of a clearly segmentable vowel plus /r/. In the weak, however, /r/ is not distinguishable in the acoustics as its own segment; rather, it influences the vowel *F*3 value. The other three speakers always realize ''for'' as weak, with an average duration of 39 ms after either stressed or unstressed syllables. Though limited, these results suggest that some speakers construct rhythmic groupings in speech production. Further work will determine whether similar effects are found in spontaneous speech.

1aSC4. A role of fundamental frequencies in the perception of emphasized words. Kyoko Nagao and Shigeaki Amano (NTT Commun. Sci. Labs., Atsugi, Kanagawa, Japan, nagao@avg.brl.ntt.co.jp)

Many previous studies have shown the role of the fundamental frequency (*F*0) on focus structure in speech production. The current study examined the role of *F*0 on focus perception. Eleven subjects of Tokyo Japanese were asked to listen to 889 *F*0-manipulated sentences and to select the word that they thought the speaker emphasized. Each stimulus included a noun phrase such as ''W1-no W2,'' where both W1 and W2 are a noun with an accent on the first mora. The *F*0 contour of each stimulus was varied by changing values of *F*0 at the three points: the *F*0 minimum point (i.e., dip) in the noun phrase and the *F*0 maximum point (peak) in each word. The result showed that *F*0 functions as a phonetic cue for a listener to perceive the emphasis on the word. We found that the relative *F*0 differences between the two peaks largely determined the subject's focus perception. Furthermore, logistic regression analysis suggests that the dip plays an important role. We will discuss the relationship between the dip and peaks for the focus perception.

1aSC5. The effects of stress and final position on sounds of Jicarilla Apache. Siri Tuttle (UCLA Dept. of Linguist., 3125 Campbell Hall, Box 951543, Los Angeles, CA 90095-1543, stuttle@ucla.edu)

The study of word-level prosody in Athabaskan languages is complicated by morphology and intonation. Stems are both frequently stressed and frequently word-final or phrase-final. This means that final lengthening confounds acoustic measurements of stress. This study of Jicarilla (Eastern) Apache attempts to sort out the acoustic effects of phrase-finality from those of stress. Syllables taken from a recorded text were measured for duration of onsets, codas and peaks. Phrasal position was noted as initial/noninitial and final/nonfinal. When only nonfinal syllables are examined, stress is marked significantly by increased vowel length ($p = 0.02$), controlling for phonemic vowel length. This distinction remains when stressed and unstressed final syllables are compared, but is weaker ($p = 0.04$). When only stressed syllables are examined, vowels are significantly longer when the syllable is final ($p = 0.0001$). However, length of codas is also significantly greater in final stressed syllables ($p = 0.027$), and this effect is not associated with stress. Thus despite the large effect of final lengthening, Jicarilla Apache still distinguishes stressed from unstressed syllables using duration within syllables. It is the ratio of rhyme elements to one another which contains the cue.

1aSC6. Stress is phonetically reset word-internally in Squamish. Linda Tamburri Watt (Univ. of British Columbia, E270-1866 Main Mall, Vancouver, BC V6T 1Z1, Canada)

The purpose of this experiment was to test a phonologically based prediction suggesting that in Squamish the domain for stress is reset word-internally in certain morphologically limited cases. This is unusual because stress systems usually assign only one primary stress within a word. This prediction was tested by first identifying the phonetic correlates of stress in this language, then applying these criteria to this morphological context. One subject was tested by recording a list of stimuli containing morphologically simplex words of varying length and words containing the morphological structure of interest. Results of this study show, first, that length, $F0$, and amplitude are all indicators of primary and secondary stress and, second, that primary stress occurs twice within a single noncompound word.

1aSC7. Sentence intonation and syllable stress in speech produced during simultaneous communication. Robert Whitehead, Brenda Whitehead (Natl. Tech. Inst. for the Deaf, Carey-1309, Rochester, NY 14623-5604), Nicholas Schiavetti, Dale Evan Metz, and Deborah Gallant (State Univ. of New York at Geneseo, Geneseo, NY 14454)

This study investigated prosodic variables of syllable stress and intonation contours in speech produced during simultaneous communication. Simultaneous communication (SC) refers to speech combined with manually coded English (sign and finger spelling) for the production of each word of an utterance and is used in communication between hearing and deaf persons. Ten normal hearing, experienced sign language users were recorded under SC and speech-only (SO) conditions speaking a set of sentences containing stressed versus unstressed versions of the same syllables and a set of sentences containing interrogative versus declarative versions of the same words. Results indicated longer sentence durations for SC than SO for all speech materials. Vowel duration and fundamental frequency differences between stressed and unstressed syllables as well as intonation contour differences between declarative and interrogative sentences were essentially the same in both SC and SO conditions. The conclusion that prosodic rules were not violated in SC is consistent with previous research indicating that temporal alterations produced by simultaneous communication do not involve violations of other temporal rules of spoken English.

1aSC8. Articulation of word and sentence stress. Patricia A. Keating, Taehong Cho, Marco Baroni (Dept. of Linguist., UCLA, Los Angeles, CA 90095-1543, keating@humnet.ucla.edu), Sven Mattys, Lynne E. Bernstein, Brian Chaney (House Ear Inst., Los Angeles, CA), and Abeer Alwan (UCLA, Los Angeles, CA)

This paper reports on the articulation of two aspects of speech prosody: word and sentence stress. Word stress is examined in reiterant renditions of minimal pairs differing in the location of main stress (e.g., DIScharge–disCHARGE), with the contrastive stress difference manifested as differences in the location of the nuclear pitch accent. Sentence stress is examined in sets of minimal sentences differing in the presence or location of the nuclear pitch accent (e.g., so TOMMY gave Debby a song from Timmy). Three talkers read these materials while movements of the jaw and tongue (using the Carstens Articulograph) or of 20 reflective infrared markers on the face (using the Qualysis Motion Capture system) were recorded. Syllables with and without contrastive pitch accents are compared along several articulatory dimensions, including: jaw displacement and duration of movement, jaw peak opening and closing velocities, jaw peak acceleration, whole head displacement, and eyebrow displacement. These comparisons allow us to determine whether the talkers articulate the focused syllables/words differently. The three talkers were selected to differ in their overall visual speech intelligibility, as rated by expert deaf speechreaders, and therefore interspeaker differences in articulation are expected. [Work supported by NSF KDI Grant 9996088.]

1aSC9. Prosodic indication of syntactic structure in quasispontaneous speech. Amy J. Schafer (Dept. of Linguist., 3125 Campbell Hall, UCLA, Los Angeles, CA 90095-1543), Paul Warren (Victoria Univ. of Wellington, New Zealand), Shari R. Speer, and S. David White (The Ohio State Univ.)

This research reports sentence production and comprehension results that demonstrate that untrained speakers reliably use prosodic structures to resolve certain syntactic ambiguities in American and New Zealand English. Speakers played a board game that required them to repeatedly produce several types of syntactically ambiguous sentences as part of the game dialogue. Speakers were not informed of the ambiguities or given instructions about how to speak. $F0$ and durational measures of critical regions and phonological transcription (ToBI) data show that speakers produced a variety of intonational structures for each syntactic structure, but consistently produced prosodic phrasal boundaries that indicated the speaker's intended resolution of the syntactic ambiguity. Prosodic disambiguation occurred even when nonprosodic disambiguating information was available in the discourse setting, and even when speakers were not consciously aware of the ambiguity. Follow-up forced-choice perception experiments on utterance fragments excised from the conversational context show that a separate group of naive, untrained listeners reliably chose the speaker's intended meaning. The results indicate that prosody–syntax correspondences are pervasive in spoken sentences, and thus should be carefully addressed in psycholinguistic models. [Work supported by NIH Grants MH51768 and DC00029, NZ/USA Cooperative Science Programme Grant CSP95/01, and Marsden Fund Grant VUW604.]

1aSC10. Right boundary strengthening in American English. Stefanie Shattuck-Hufnagel (Speech Group, RLE, 36-511 MIT, 77 Massachusetts Ave., Cambridge, MA 02139, stef@speech.mit.edu)

A prominent theme in recent phonetic research is the role of constituent structure in governing the phonetic implementation of phonemic segments. For example, Fougeron and Keating (1997) describe an articulatory strengthening effect for some constituent-initial consonants which reflects the prosodic constituent hierarchy: utterance-initial /n/ is produced with greater tongue contact than phrase-initial /n/, which is in turn produced with more contact than word-initial /n/, etc. This articulatory strengthening is reflected to some extent in acoustic measures of duration. Is such strengthening limited to the left edge of constituents, or can it also occur at

the right edge? Results from read laboratory speech show that the occurrence of aspiration in a word-final /k/ in American English is governed in part by the nature of the prosodic boundary following the word. Results for five speakers showed that (a) utterance-final /k/, as in ''Please say bake,'' was almost always aspirated (as determined by acoustic measures of the wave form); (b) word-final /k/ before a word-initial schwa, as in ''Please say bake again,'' was almost never aspirated, and (c) phrase-final /k/ before word-initial schwa, as in ''Please say bake, again,'' was sometimes aspirated and sometimes not. [Work supported by NIH DC02125-05.]

1aSC11. Accentual lengthening in Mandarin Chinese. Yiya Chen (Dept. of Linguist., SUNY at Stony Brook, Stony Brook, NY 11794-4376, yiyachen@yahoo.com)

This study investigated the effects of accentual lengthening under contrastive focus in Beijing Mandarin, a tone language, as a comparison to the results of similar studies in stress-accent languages. Two questions were asked. First, what is the domain of lengthening? Second, what are the structural influences on accent lengthening? Four different renditions of each test sentence were elicited from the subjects with contrastive focus on four different words to convey different pragmatic meanings. The results suggest: (1) the domain of accentual lengthening is primarily the focused word, with weak spread to the following word; (2) the spread of lengthening is sensitive to the morphosyntactic/prosodic relation of the focused word and the following word; and (3) while both onset and nucleus of the focused syllable exhibit lengthening, the spreading effect of lengthening is observed more reliably on nuclei than on onsets. These results agree to a large extent with studies on accentual lengthening in English and Dutch [A. Turk and L. White, J. Phonet. **27**, 171–206 (1999); T. Cambier-Langeveld and A. Turk, J. Phonet. **27**, 255–280 (1999)]. The similarity of accentual lengthening in Mandarin and stress-accent languages suggests that some effects of accentual lengthening may be universal. [Work supported by NSF No. SBR 9600930.]

1aSC12. Post obstruent tensing in Korean: Its status and domain of application. Sahyang Kim (Dept. of Linguist., UCLA, 3125 Campbell Hall, Box 951543, Los Angeles, CA 90095-1543, sahyang@ucla.edu)

The criteria that distinguish the phonological rules from the phonetic rules based on categoricality versus gradiency have drawn considerable attention (cf. Keating, 1990; Pierrehumbert, 1990). Korean Post Obstruent Tensing (POT) rule has been claimed, based on acoustic data, to be a phonological rule and that its domain of application is an accentual phrase (AP) (e.g., Jun, 1998). This paper questions the status of POT and its domain of application based on articulatory and acoustic data. By POT, a syllable initial lenis consonant (e.g., /t/, /s/) becomes tense (e.g., /t*/, /s*/, respectively) after an obstruent coda. It is known that lenis obstruents have a longer VOT, shorter closure durations, and shorter linguopalatal contact areas than tense obstruents. To investigate if POT is really a categorical process and if its application is definable in terms of an AP, acoustic and articulatory (i.e., EPG) data of derived and underlying tense obstruents were examined in four different prosodic positions: utterance-initial, intonational phrase-initial, AP-initial, AP-medial. Sentences containing a target sequence, /VpCV/, where C=/t/, /t*/, /s/, /s*/, were produced by two Korean speakers, at normal and fast speech rates. The results will shed light on the status of POT in the grammar of Korean.

1aSC13. Prosodic domain effects on noninitial syllables. Marie Huffman (Dept. of Linguist., SUNY at Stony Brook, Stony Brook, NY 11794-4376) and Jennifer Hataier (Southern Connecticut State Univ., New Haven, CT)

Previous literature has documented strengthening of segments adjacent to prosodic domain boundaries. This strengthening might be expected to affect only the segment next to the boundary. Yet, Fougeron and Keating [J. Acoust. Soc. Am. **101**, 3728–3740 (1997)] found that some noninitial vowels showed strengthening in a domain-initial syllable, as did some onset consonants in final syllables. To test whether effects of prosodic domain boundaries extend beyond the domain-adjacent syllable, consonant and vowel durations were measured for the second, stressed, syllable in the place names Apollo, Atami, and Acosta, comparing utterance-, intonation phrase- and word-initial position. Results for four female and one male speaker indicate that while there is rarely any effect on the consonant, the stressed vowel in the nonboundary adjacent syllable shows domain effects in more than half of the items tested, with longer durations in higher domains, particularly in Atami. These surprising nonlocal effects suggest that: (1) the place of articulation effect may be due to simple distance from the boundary, since [t]s were usually shorter than the other consonants; and (2) domain edge strengthening may involve a change in articulatory setting, the onset of which is controlled more carefully than the offset. [Work supported by NSF SBR9600930.]

1aSC14. An articulatory study of prosodically conditioned V-to-V coarticulation in English. Taehong Cho (Dept. of Linguist., UCLA, 405 Hilgard Ave., Los Angeles, CA 90095)

Recent studies have indicated that vowels in prosodically strong positions (e.g., in stressed syllables and at edges of prosodic boundaries) are not only strongly articulated, but also resistant to coarticulation with neighboring vowels. This paper further examines vowel-to-vowel coarticulation in English by analyzing extensive articulatory data from six American English speakers, using the Electromagnetic Articulograph (EMA). It is hypothesized that vowels in prosodically strong positions are more resistant to coarticulation with their neighbors, and at the same time encroach more on their neighbors. To test this, sentences were designed so that they included /V1#bV2/ where V1 and V2 were manipulated, resulting in /a-a/, /i-i/ (control condition) and /a-i/, /i-a/ (test condition). Vowels also varied in sentence stress (accented versus unaccented) and in the intervening boundaries (# = Word, ip, IP). The vertical and horizontal positions of three tongue points and jaw are examined at five different points (onset, first quarter, middle, three quarters and end) in the vowel, to assess how much of the vowel articulation is anticipated or carried over at different points of the vowel. This shows variation in degree of V-to-V coarticulation under various prosodic conditions. [Work supported by NSF doctoral research grant.]

1aSC15. Relationships among temporal patterns of speech production in English. Bruce L. Smith (Dept. of Commun. Sci. and Disord., Northwestern Univ., 2299 N. Campus Dr., Evanston, IL 60208)

Although various temporal patterns have been noted in English, possible relationships among these different characteristics generally have not been examined. In part, this may be because it is assumed that speakers show temporal patterns such as phrase-final vowel lengthening, vowel lengthening preceding voiced obstruents, etc., to essentially the same degree. However, recent findings suggest that there can be considerable variation among speakers in the extent to which they manifest such temporal properties. Thus, the present study investigated the possibility that subjects who exhibit a particular temporal property in their speech to a greater/lesser degree than other speakers might also tend to show other patterns to a greater (or lesser) extent. The same temporal characteristics were also examined to determine whether they varied as a result of speakers different syllable durations. A number of relationships among the various parameters were observed. For example, speakers with shorter syllable durations tended to show greater phrase-final vowel lengthening than speakers with longer syllable durations. In contrast, subjects with shorter syllable durations showed less vowel lengthening preceding voiced versus voiceless consonants than speakers with longer syllable durations. Possible reasons for these temporal variations and interrelationships will be considered.

1aSC16. Prosodic prominence on negation in various ''registers'' of U.S. English. Malcah Yaeger-Dror, Sharon K. Deckert, and Lauren Hall-Lew (Dept. of Cognit. Sci., Univ. of Arizona, Tucson, AZ 85721)

Since negatives express cognitively critical information, researchers have concluded that negatives ''should'' be prosodically prominent. Acoustic analysis of pitch tracks from isolated sentences [O'Shaughnessy and Allen (1982)] and the BU Radio News corpus [Hirschberg (1990)] support this theoretical claim. This study presents evidence of prosodic variation on ''not''-negatives used in a range of different social situations distinguished along Biber's ''register'' continuum from informative to interactive. The paper will show that both the relative quantity and the type of intonational prominence varies in different registers. Informative data discussed will include both LDC's BU Radio News and the Air Traffic Control corpus. Interactive corpora include both the Switchboard corpus of polite conversations between strangers and archived presidential debates. The percentage of prominent tokens is quite low in Switchboard, higher in debates, and much higher in the informative corpora. ToBI's H* category is dominant only in the most purely informative settings, with other contours preferred in interactive registers. [Work supported by NSF.]

MONDAY AFTERNOON, 4 DECEMBER 2000 SCHOONER/SLOOP ROOMS, 1:20 TO 5:00 P.M.

Session 1pAO

Acoustical Oceanography: Bioacoustics II

Timothy K. Stanton, Chair
Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Bigelow 201, Woods Hole, Massachusetts 02543-1053

Chair's Introduction—1:20

Invited Paper

1:30

1pAO1. A mini-tutorial on zooplankton acoustics. Peter H. Wiebe (Woods Hole Oceanogr. Inst., Redfield Bldg., Woods Hole, MA 02543)

The techniques for use of high-frequency acoustical systems for the study of zooplankton distribution, biomass, abundance, size distribution, and behavior have been the focus of intense effort and development during the past two decades. Significant theoretical and experimental progress has been made in understanding how these animals scatter sound. A review of the major aspects of this problem will be presented, including current field survey techniques, instrumentation, laboratory measurements, and scattering models. Issues such as confounding factors in data interpretation and determination of model parameters will be discussed.

Contributed Papers

2:30

1pAO2. Three-dimensional acoustic scattering models for elongated fluid-like zooplankton. Andone C. Lavery, Dezhang Chu, Duncan E. McGehee, and Timothy K. Stanton (Woods Hole Oceanogr. Inst., Appl. Ocean Phys. and Eng., Woods Hole, MA 02543)

Accurate acoustic scattering models that correctly incorporate organism size, shape, material properties, and orientation are a critical component for the remote classification and detection of marine organisms, such as fish and zooplankton. Zooplankton that fall into the elongated fluid-like category are of particular importance, primarily due to their high natural abundances. In order to obtain a better understanding of the scattering properties of this class of organism, extensive laboratory measurements of the frequency (50 kHz to 1 MHz) and angular (all angles in two planes with 1-deg resolution) characteristics of the acoustic backscattering from live individual decapod shrimp (*Palaemonetes vulgaris*) have been performed. An acoustic scattering model based on the distorted wave Born approximation (DWBA) has been developed for these elongated fluid-like organisms. This model makes use of progressively more realistic and complex representations of the animal shape. The most sophisticated of these representations involves fully three-dimensional digitizations of the animal exterior, obtained using high-resolution computed tomography (CT). A primary focus of this study is to determine the practical conditions under which it is necessary to make use of such high-resolution, computer-intensive digitizations of animal shape.

2:45

1pAO3. Multistatic multifrequency scattering from zooplankton. Charles F. Greenlaw, D. Van Holliday, and Duncan E. McGehee (Bae Systems, 4669 Murphy Canyon Rd., San Diego, CA 92123)

Multifrequency backscattering measurements can be inverted to estimate size abundance of zooplankton if a suitable model for backscattering is known. Physics-based models generally include the physical properties (density and compressibility) as input parameters. Computer manipulations of one such zooplankton scattering model suggest that physical properties as well as size might be estimated from simultaneous multistatic and multifrequency scattering measurements. Preliminary results from laboratory measurements of silicone rubber targets will be presented and extraction of physical properties discussed. [Work supported by ONR Code 322 BC.]

3:00

1pAO4. Seasonal variations of density and sound-speed contrasts of the krill *Euphausia pacifica*. Tohru Mukai, Kohji Iida, Hisayuki Mikami, and Ryuichi Matsukura (Grad. School of Fisheries Sci., Hokkaido Univ., 3-1-1 Minato-cho, Hakodate, Hokkaido, 041-8611 Japan)

In recent years, acoustic technology has been used extensively to estimate krill abundance. Acoustic assessment is required to gain a precise estimate of krill target strength. However, predictions of target strength from theoretical scattering models are often influenced by the swimming

angles and density of the krill, and sound-speed contrasts between krill and seawater. Density and sound-speed contrasts are known to show annual cycles. In this study, seasonal variations of the specific density and sound-speed contrasts of *Euphausia pacifica* are presented. Biological sampling was carried out during twilight, when the sound-scattering layer migrates up to the surface. The specific densities of *E. pacifica* were measured, using a series of variable density glycerol solutions, within 48 h of net sampling. Sound-speed measurements were performed 2 hours after net sampling, using a T-shaped velocimeter with two transducers mounted on the ends of a horizontal tube. In 1999, the seasonal changes in specific densities and sound speed contrasts were ca. 1% and 3%, respectively. When calculated using Stanton's straight cylinder model, these seasonal changes yielded a difference of approximately 5 dB in the target strength of *E. pacifica*.

3:15–3:30 Break

3:30

1pAO5. Broadband echo spectra from euphausiids and copepods. Kenneth G. Foote (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, kfoote@whoi.edu), Tor Knutsen (Inst. of Marine Res., N-5024 Bergen, Norway), Philip R. Atkins, Claire Bongiovanni, David T. I. Francis (Univ. of Birmingham, Birmingham B15 2TT, UK), Peter K. Eriksen, Mette Torp Larsen, and Tom Mortensen (RESON A/S, DK-3550 Slangerup, Denmark)

A seven-octave-bandwidth echo sounding system [J. Acoust. Soc. Am. **105**, 994 (1999)] was used to observe diverse zooplankton along the coast of western Norway during the period 28 April–9 May 1999. Echo spectra were obtained from freshly caught, swimming specimens of the euphausiid *Meganyctiphanes norvegica* and copepod *Calanus finmarchicus ex situ* in a tank mounted on the stern deck of R/V JOHAN HJORT. Echo spectra were also obtained from scatterers *in situ* at 15-m depth in the same sea areas where the experimental animals were caught. Pulse compression was used to reduce the risk of analyzing overlapping echoes. All echo spectra were absolute, as the system was calibrated by the standard-target method using a 10-mm-diameter sphere of tungsten carbide with 6% cobalt binder. Comparison of echo spectra over the nominal bandwidth 1600–2600 kHz supports the claim of acoustic classification. Additional evidence is provided by reference to modeling computations [J. Acoust. Soc. Am. **105**, 1050 (1999); **105**, 1111 (1999)] where the bodies were assumed to be weakly scattering and fluid-like, with representative morphometries and assumed or measured properties of mass density and longitudinal-wave sound speed. [Work supported by the EU through RTD Contract No. MAS3-CT95-0031 during the project.]

3:45

1pAO6. *In situ* target tracking of individual krill (*Meganyctiphanes norvegica*) in the Oslo fjord. Anders Røstad (Oregon State Univ., Hatfield Marine Sci. Ctr., 2030 SE Marine Science Dr., Newport, OR 97365, anders.roestad@noaa.gov) and Stein Kaartvedt (Univ. of Oslo, Blindern, N-0316 Oslo, Norway)

Krill (*Meganyctiphanes norvegica*) were studied during and after their diel vertical migration to near-surface water using a Simrad EK500 echosounder with a split-beam 120-kHz hull mounted transducer. Target tracking of individual krill provided information on target strength, swimming velocity, and 3-D swimming patterns through the acoustical beam. Both prospects and limitations of the split-beam target tracking technique in behavioral studies are discussed.

4:00

1pAO7. Three-dimensional acoustic tracking of krill with a multibeam sonar. Alex De Robertis, Chad Schell, and Jules S. Jaffe (Scripps Inst. of Oceanogr., La Jolla, CA 92093, aderobertis@ucsd.edu)

The swimming behavior of individual zooplankton mediates how the animals experience their spatially heterogeneous environment, and consequently, has an important effect on population dynamics. However, current understanding of zooplankton swimming behavior is limited by conventional zooplankton sampling techniques. Here, we describe the use of a high-resolution multibeam sonar [Jaffe *et al.*, Deep Sea Res. **42**, 1495–1512 (1995)] to reconstruct the swimming trajectories of individual krill (*Euphausia pacifica*) in Saanich Inlet, British Columbia. The instrument was deployed in the deep scattering layer at depth during the day, and in near-surface strata during twilight periods of vertical migration and at night. Successive acoustically determined animal positions from these records are linked using a simple target tracking algorithm resulting in $>10^4$ trajectories of several seconds duration. The spatial positions of the tracked targets are improved by applying an algorithm [J. S. Jaffe, J. Acoust. Soc. Am. **105**, 3168–3176 (1999)] developed to localize targets within the acoustic beams. These methods permit quantitative analyses of swimming behavior of undisturbed krill in their natural environment for the first time.

4:15

1pAO8. Experimental verification of an algorithm for animal localization using a multibeam sonar system. Jules S. Jaffe, Chad Schell, and Alex De Robertis (Scripps Inst. of Oceanogr., La Jolla, CA 92093)

Behavioral inferences of zooplankton activity from data generated with active multibeam sonar systems require high precision in localization due to the small extent of the animal movements. If target localization is limited to whether a target is within one beam or another, sonars with many beams and high complexity are required. Sonars which have multiple beams with measurable side lobes have the capability of achieving adequate resolution in situations of high signal to noise. In Jaffe [J. Acoust. Soc. Am. **105**, 3168–3175 (1999)], a maximum likelihood estimator for position was proposed. The algorithm utilizes a minimum mean square estimator in amplitude in order to estimate target position to a resolution smaller than the sonar's beamwidth. Experimental verification of the application of this algorithm to data aquired in a test tank with simultaneous verification of position with an optical imaging system is presented. The methodology employed uses a pair of cameras in order to simultaneously estimate the three-dimensional location of a target which is simultaneously being observed with a three-dimensional multibeam acoustical system [Jaffe *et al.*, Deep Sea Res. **42**, 1495–1512 (1995)]. The results indicate that the algorithm can produce an accurate estimate of three-dimensional target positional information providing accurate calibration is performed.

4:30–5:00

Panel Discussion

Session 1pBB

Biomedical Ultrasound/Bioresponse to Vibration: Topical Meeting: Physics of Echo-Contrast Agents II

E. Carr Everbach, Chair
Department of Engineering, Swarthmore College, 500 College Avenue, Swarthmore, Pennsylvania 19081-1397

Invited Papers

1:30

1pBB1. Ultrasound bioeffects associated with contrast agents. Diane Dalecki (Dept. of Elec. and Computer Eng., Univ. of Rochester, Rochester, NY 14627, dalecki@ece.rochester.edu)

Ultrasound contrast agents are stabilized microbubbles used to enhance diagnostic imaging. Various studies indicate that certain bioeffects can result from the interaction of ultrasound with tissues containing contrast agents. The presence of ultrasound contrast agents in blood has been shown to increase the extent of hemolysis resulting from exposure to pulsed ultrasound *in vitro* and *in vivo*. The presence of contrast agents in tissues also increases the susceptibility of many tissues to damage from pulsed ultrasound or lithotripter fields. For example, injection of Albunex during exposure of mice to a piezoelectric lithotripter field with amplitude of only 2 MPa produced hemorrhages in most soft tissues, such as fat, muscle, kidney, stomach, and bladder. Contrast agents in the vasculature can also decrease the threshold for the production of premature cardiac contractions resulting from exposure to pulses of ultrasound. This lecture will provide an overview of some bioeffects of ultrasound in tissues containing ultrasound contrast agents.

1:45–2:45

Panel Discussion on Bioeffects of Echo-Contrast Agents

2:45–3:00 Break

3:00

1pBB2. Sonoporation and ultrasound transfection assisted by contrast agents. Junru Wu (Dept. of Phys., Univ. of Vermont, Burlington, VT 05405, jwu@zoo.uvm.edu)

Gene therapy provides potentially effective treatment of a wide variety of diseases. Generally speaking, all gene therapy protocols require introduction of foreign nucleic acids, most often DNA, into the target cells and tissue, where gene expression is induced. However, low permeability of tissue and cells causes efficient barriers to entry of foreign DNA. Therefore, gene therapy requires transfection techniques that promote the uptake and expression of foreign DNA by a cell. There are two types of transfection techniques—viral and nonviral. Among the nonviral transfection methods, sonoporation and ultrasound-induced transfection are relatively new methods. Recent *in vitro* studies [Bao *et al.*, Ultrasound Med. Biol. **23**, 953–959 (1997); Greenleaf *et al.*, *ibid.* **24**, 587–595 (1998); Ward *et al.*, J. Acoust. Soc. Am. **105**, 2951–2957 (1999)] have shown that the presence of contrast agents dramatically enhances the efficiency of the transfection and also reduces the required acoustic pressure. This presentation will focus on reviewing the recent publications of sonoporation and ultrasound transfection using contrast agents. The possible physical mechanisms involved and main advantages of this technique will be discussed.

3:15–4:15

Panel Discussion on Novel Uses of Echo-Contrast Agents

4:15–4:30 Break

4:30

1pBB3. Future directions in ultrasound contrast agent research. Charles C. Church (Acusphere, Inc., 38 Sidney St., Cambridge, MA 02139, church@acusphere.com)

Since the introduction of Levovist® in Europe and Albunex® in the United States, interest in ultrasound contrast agents has increased enormously. This change is reflected in the ensuing rapid pace of technical development, and the depth of the discussions during this topical meeting. Before Albunex®, it was essentially impossible to convince an equipment manufacturer to devote time to the development of contrast-specific imaging modalities. After Albunex®, harmonic gray-scale imaging was in the clinic within the year, followed by harmonic color Doppler imaging, power Doppler, pulse inversion gray scale, pulse inversion Doppler, power modulation, coherent contrast imaging, and so on. The trend continues today as old techniques are improved, new techniques (e.g., subharmonic imaging) are tested and new approaches (e.g., phase-shift imaging) are conceived. Similarly, rapid advances have occurred in other areas, with models now including the details of shell viscoelasticity, *in vitro* studies now imaging single bubble

dynamics to microsecond resolution, and specially designed microbubbles being used for targeted imaging, drug delivery, and gene therapy. Even bioeffects research has ''advanced'' with the discovery of imaging-modality-dependent contrast agent effects such as induction of premature ventricular contractions. Selected aspects of each of these areas will be discussed.

4:45–5:45

Panel Discussion on Modeling of Echo-Contrast Agents

MONDAY AFTERNOON, 4 DECEMBER 2000 PACIFIC SALON D, 1:20 TO 2:20 P.M.

Session 1pNSa

NOISE-CON and Noise: Plenary Session—Product Sound Quality From Perception to Design

John J. Van Houten, Chair

J. J. Van Houten and Associates, 3320 East Chapman Avenue, #323, Orange, California 92869-3811

1:20

1pNSa1. Product sound quality: From perception to design. Richard H. Lyon (RH Lyon Corp, 691 Concord Ave., Cambridge, MA 02138)

This discussion of product sound quality is concerned with the relation between the work of product designers and the perceptions of consumers/users regarding the acceptability of the sound. Designers make choices regarding structure, materials, and components in a product. The tools they need should allow them to anticipate the effect of those choices on sound quality. This presentation recounts the role of psychoacoustics in product design and product acceptability and notes the results of that work in metrics for sound quality and consumer/user perceptions about the product. The successes and drawbacks of this activity are noted. Recent work on a new paradigm, using sensory profiles for sound as an intermediary between metrics and perception, is described along with results using this procedure. Future developments of the procedure are outlined. [Work supported in part by National Science Foundation Dynamic Systems and Controls Program and by RH Lyon Corp .]

MONDAY AFTERNOON, 4 DECEMBER 2000 PACIFIC SALON F, 2:35 TO 5:10 P.M.

Session 1pNSb

Noise, NOISE-CON and Engineering Acoustics: Performance Assessment of Acoustical Test Rooms

Kenneth A. Cunefare, Chair

School of Mechanical Engineering, Georgia Institute of Technology, Graduate Box 286, Atlanta, Georgia 30332-0405

Chair's Introduction—2:35

Invited Papers

2:40

1pNSb1. Qualifying a reverberation room for sound absorption measurements. Alf Warnock (Inst. for Res. in Construction, Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

ASTM C423, the standard method for measuring sound absorption in a reverberation room, was recently revised extensively and now includes tests to verify that the room has acceptable acoustical properties. Further changes are planned by the ASTM task group. This paper will present sets of measurements made in the reverberation rooms at the National Research Council Canada illustrating the effects of some physical factors, such as number of loudspeakers, number of microphones, the need for rotating diffusers, and the number of specimen positions that should be used for different specimens. Practical problems facing laboratory operators will be discussed. The precision required for measurement in a room will be determined to some extent by the reproducibility of the test method. There is little point in reducing the precision of the measurement in individual rooms once it is significantly less than the test reproducibility. Further changes that might be made to C423 will be presented.

3:00

1pNSb2. Dynamic insertion loss measurements of sound attenuators at low frequencies. Jerry G. Lilly (JGL Acoust., Inc., 5266 NW Village Park Dr., Issaquah, WA 98027) and James D. Partain (PCI Industries, Inc., Fort Worth, TX 76111)

For decades the duct silencer industry has relied on ASTM E 477 as a standard for measuring the acoustic and aerodynamic properties of passive duct silencers. Historically, there has been significant concern regarding the accuracy and reliability of the measured dynamic insertion loss values for duct silencers at 125 Hz and below. The uncertainty at low frequencies has been blamed on several factors including: (1) inaccurate determination of sound level in reverberation chambers due to inaccurate spatial averaging (e.g., low modal density), and (2) varying acoustic output power from the electronic signal source due to a different impedance seen by the loudspeaker (with and without the test specimen in the duct). This paper presents a third reason for the low-frequency variability, which is likely to be more significant than all others combined. It will show that the problem is caused by acoustic resonances in the empty test duct, and these resonances will occur in every facility that is constructed in accordance with the test standard. The authors will present test data collected in a NVLAP accredited laboratory verifying this low-frequency effect, and also recommend specific changes to ASTM E477 which could eliminate this problem.

3:20

1pNSb3. Anechoic wedge design and development/anechoic chamber qualification testing. Alan Eckel (Eckel Industries, 155 Fawcett St., Cambridge, MA 02138, ae@eckelacoustic.com) and Istvan Ver (Eckel Industries, Stowe, MA 01775)

The initial research conducted at Harvard University's Electro Acoustic Laboratory for the development of the anechoic chamber will be reviewed in this presentation. These findings, which were originally presented in the Office of Scientific Research and Development National Defense Research Committee Report No. 4190 (OSRD, No. 4190), will be discussed. The research conducted by Leo L. Beranek and his colleagues during WWII established the basic design and construction for anechoic wedges and chambers for years to come. Contemporary design and materials used in anechoic wedge construction have evolved to encompass a variety of room acoustic treatments, to include fiberglass, foam, and metallic wedges. Design and construction of current commercially available chambers and wedges will be reviewed. The second half of the presentation will focus on the development of anechoic and hemi-anechoic chamber qualification procedures to current ISO and ANSI standards. The use of various sound sources will be discussed as well as the correlation between these standards and other test methods in determining chamber performance.

3:40

1pNSb4. Experiences with anechoic chamber qualification per ISO 3745 and ASA/ANSI S12.35. Kenneth A. Cunefare (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405)

ISO 3745 and ASA/ANSI S12.35 contain recommended procedures for the performance qualification of anechoic and semianechoic rooms. This paper presents experience with a number of different methodologies and analyses used in a series of tests configured to comply with these specifications. The methods include discrete measurements at fixed separation distances with multiple microphones, measurements taken with a single microphone repositioned between each measurement, and a computer-controlled continuous traverse system. Experience with different instrumentation chains will be briefly addressed. Also, a comparison of test results using pure tone and pink noise excitation will be presented. Finally, an algorithm will be described which, for a given continuous traverse test data set, determines the best acoustic source center and source offset distance to minimize the deviation from the permissible bounds contained within the specifications.

4:00

1pNSb5. On sound power determination: The uncertainty in sound power determined in a reverberation room based on a round-robin. Richard J. Peppin (Scantek, Inc., 916 Gist Ave., Silver Spring, MD 20910, peppinr@asme.org)

In this paper I will present the results of an extensive round-robin sponsored under the NIST National Voluntary Laboratory Accreditation Program for Acoustics. The round-robin tested the interlaboratory variability of sound power measured using the reverberation room sound power standards ANSI S12.31-33 and ISO 3741-3743. Also I will discuss the ramifications of uncertainty on sound-pressure estimation.

4:20

1pNSb6. Research into quality assessment methods for anechoic chambers. Francis Babineau, Jr. and Brandon Tinianow (Johns Manville, 10100 W. Ute Ave., Littleton, CO 80127, babineau@jm.com)

The goal in qualification testing for any anechoic or hemi-anechoic chamber is to define a volume within which acoustic measurements can be performed accurately in a free-field condition. In this study, both commonly accepted and novel techniques are employed to assess the acoustical performance of a hemi-anechoic chamber and ultimately its relevancy for qualified testing. Two recognized international test methods are discussed, as well as various alternate methods. The alternate methods include variations on the existing standards, utilizing acoustic intensity measurements and using an alternate sound source. The alternate sound source more closely approaches actual sound power testing conditions, in that it does not have a well-defined acoustic center. Considerations for performance assessment include impact on sound power measurements, ease of assessment, and comparisons of results to those of accepted methods. Recommendations for improvements to contemporary assessment techniques as well as future work will also be presented.

4:40

1pNSb7. Design and construction of a convertible hemi/anechoic acoustical laboratory for testing space flight hardware at the NASA Glenn Research Center. Beth A. Cooper (NASA John H. Glenn Res. Ctr. at Lewis Field, 21000 Brookpark Rd., M.S. 86-10, Cleveland, OH 44135, beth.a.cooper@grc.nasa.gov)

At the NASA Glenn Research Center, science experiment payloads developed for the International Space Station are designed to meet noise emission criteria that promote hearing conservation and speech intelligibility goals. A key component in the successful design of low-noise hardware is the frequent use of acoustical testing as a diagnostic and design verification tool. NASA has designed and constructed a convertible hemi/anechoic acoustical testing laboratory that provides convenient access to accurate and repeatable sound pressure level measurements as well as state-of-the art capabilities for sound power level testing per ANSI S12.35 using a multichannel PC-based data acquisition system. The laboratory consists of a 100-Hz hemianechoic chamber, vibration-isolated to 3 Hz, with $21' \times 17' \times 17'$(h) interior working dimensions and removable floor wedges. A separate, sound-isolated control room doubles as a test support equipment enclosure when testing flight hardware that requires remote connections to high-noise support equipment and utilities. Movable modular furnishings facilitate reconfiguration of the enclosure, and data acquisition and test control functions are easily relocated to an adjacent quiet area. The Glenn Research Center Acoustical Testing Laboratory also offers a full range of noise control services to assist payload developers with meeting the ISS noise emission criteria.

4:55

1pNSb8. Noise and vibration isolation design for a vibration testing laboratory. David A. Nelson (Nelson Acoustics Engineering, Inc., P.O. Box 879, Elgin, TX 78621) and Thomas Stewart (Cisco Systems, Inc., San Jose, CA 95134-1706)

Cisco Systems new Engineering Approvals Center provides an example of the successful application of noise and vibration control principles in a mixed-use building. The Vibration Laboratory portion of the building is designed for maximum throughput and comprises a number of powerful noise and vibration sources, including four hydraulic shaker tables with hydraulic pumps, an electrodynamic shaker table, two drop shock tables, and a package drop area. Because office spaces, conference facilities and sensitive high-performance EMC chambers are located directly above and adjacent to the Vibration Laboratory, a considerable degree of isolation was designed into the structure to preserve an officelike environment in these areas. Design and installed performance of critical building elements will be discussed, including the isolated slab and vibration breaks, noise-isolating control room, walls, windows, doors and ceiling.

MONDAY AFTERNOON, 4 DECEMBER 2000 PACIFIC SALON D, 2:35 TO 5:05 P.M.

Session 1pNSc

Noise and NOISE-CON: Noise Standards—Challenges to Quieter Products

Robert D. Hellweg, Jr., Chair
Compaq Computer Corporation, MR 01-3/D3, 200 Forest Street, Marlborough, Massachusetts 01752

Chair's Introduction—2:35

Invited Papers

2:40

1pNSc1. Noise labeling for consumer products. Paul Schomer (USA-CERL-CNN, P.O. Box 9005, Champaign, IL 61826-9005, schomer@uiuc.edu)

Currently, ISO 4871-1996 (Acoustics–Declaration and verification of noise emission values of machinery and equipment) uses A-weighted or spectral sound power levels as metrics of choice for noise labeling. Also included is a measure of the likely variations in levels that result from machine to machine variability, measurement to measurement variability, and site to site variability. This labeling scheme is adequate to label machinery for use in commercial or industrial settings. Here, plant engineers or audiologists, etc., can readily interpret and use data. But a two-number label using the sound power level in decibels is not optimum for conveying information to the consumer. This paper suggests an alternate form for the data. Sound power is still the physical quantity measured, and, as suggested by others, a one-number system is used. The statistical variation is included by reporting the value that will not be exceeded in 19 out of 20 measurements. But the metric reported is one that is roughly proportional to judgments of loudness, and, as such, should provide greater meaning and clarity to the user. Hopefully, with comprehensible noise labeling, the consumer will be able to include noise in his/her choice of products.

3:00

1pNSc2. The cylindrical microphone array: A proposal for use in international standards on sound power level measurement. Matthew A. Nobile, Jennifer A. Shaw (IBM Hudson Valley Acoust. Lab., Bldg. 704, M.S. P226, 2455 South Rd., Poughkeepsie, NY 12603), and Brian Donald (9 Creek Bend Rd., Poughkeepsie, NY 12603)

To determine the sound power level of a stationary noise source, a hypothetical measurement surface is defined which envelops the source under test. The sound pressure level is measured at selected microphone positions on this measurement surface, and the sound power level computed. A previous paper [M. A. Nobile and J. A. Shaw, Inter-Noise **99**, 1535–1540] introduced the concepts of the cylindrical microphone array and cylindrical measurement surface and presented results from initial measurements. The array

was suggested as a more practical alternative to the current parallelepiped or hemispherical arrays for measuring the sound power level of sound sources in hemi-anechoic chambers according to standards such as ISO 3744, or the Information Technology and Telecommunications industry test code ISO 7779. It was also suggested that for certain types of machines, such as tall, rectangular computer frames, or other equipment with similar aspect ratios, the cylindrical measurement surface may yield results that are more accurate. In this paper, the results of further experiments and measurements using the cylindrical array will be presented, along with a proposal for incorporating this new array into the international standards on sound power level measurements.

3:20

1pNSc3. Method for evaluating the attenuation of acoustical covers used in high-end computer systems. Jennifer A. Shaw, Matthew A. Nobile (IBM Hudson Valley Acoust. Lab., Bldg. 704, M.S. P226, 2455 South Rd., Poughkeepsie, NY 12601), and Brian Donald (Poughkeepsie, NY 12603)

High-end computer systems, such as those used in the Information Technology and Telecommunications (ITT) industry, have become increasingly noisy in recent years, due to the amount of airflow needed to keep them functioning properly. As sound power levels increase, the systems run the risk of exceeding their noise-limit specifications. One solution is to incorporate acoustical covers as a noise control measure. By attenuating the amount of noise emitted from the machine, the overall measured sound power level will decrease, thereby allowing the system to pass specifications. However, a method for evaluating the amount of attenuation these acoustical covers provide is needed. To date, the attenuation has been measured by using actual system frames that are under development, which are often hard to acquire and introduce too many variables into the measurement. Also, both front and rear covers are needed for the measurements, so that the attenuation of just one cover is not easily obtained. This paper proposes the use of a test box as an evaluation tool for measuring the attenuation of a single cover. The test box provides a uniform and repeatable method that yields a more useful metric for rating cover performance.

3:40

1pNSc4. A comparison of two methods for the evaluation of prominent discrete tones: Phase 3. Anne C. Balant (State Univ. of New York at New Paltz, HUM 18, Ste. 6, New Paltz, NY 12561 and IBM Hudson Valley Acoust. Lab., balant@us.ibm.com), Robert D. Hellweg, Jr. (Compaq Computer Corp., Marlboro, MA 01752), and Matthew Nobile (IBM Hudson Valley Acoust. Lab., Poughkeepsie, NY 12601-5400)

This paper presents results of Phase 3 of the study of prominent discrete tones by the Inter Committee Working Group (ICWG) from Information and Technology Equipment (ITTE). Progress and earlier results were reported at INTER-NOISE 99, the 139th Meeting of the Acoustical Society of America, and INTER-NOISE 2000. Phase 3 is a round robin involving 40 signals (some from ITTE products and some artificial sounds), which were recorded and distributed to participants as .wav files. Each of several ITTE company laboratories determined the objective tone-to-noise ratios and prominence ratios of the signals and/or provided subjective ratings. The paper presents comparisons of the objective and subjective ratings from the different laboratories. The relative success of the two objective methods in predicting the subjective ratings is discussed, as are the underlying issues and recommendations of the task group.

4:00

1pNSc5. Computer sound quality: Masking, loudness and prominence. David A. Nelson (Nelson Acoustics Engineering, Inc., P.O. Box 879, Elgin, TX 78621)

Computers and business equipment were one of the first applications for standardized sound quality metrics. Various product specifications (emanating primarily from Europe) defined criteria for A-weighted sound pressure and sound power levels as well as the absence of prominent tones and impulsive noise. Standards such as ISO 7779, ANSI S12.10 and ECMA 74 have provided the means for assessing these quantities. Of particular interest at the moment is the prominent tone, for which there are two accepted metrics: tone-to-noise ratio and prominence ratio. Experience has shown that both metrics occasionally fail to track listeners perception of the prominence of tones or tonal complexes. An alternate prominence analysis is constructed based on psychoacoustical factors, using a method proposed by Zwicker for identifying the loudness of any sound masked by another. The accepted metrics are compared and contrasted with the masked loudness method using carefully selected test cases.

Contributed Papers

4:20

1pNSc6. Sound quality of contemporary personal-computers. Ya-Wen Chou and Ming-San Cheng (D300, Bldg. 64, No. 195-6, Sec. 4, Chung Hsing Rd., Chutung Hsinchu, Taiwan 310, ROC, gloria@itri.org.tw)

Development of sound quality analysis in product design and marketing has drawn much attention in recent years. Many makers nowadays would like to know their products sound quality and to improve them instead of just reducing the overall sound pressure level. The noise of personal-computers is characterized as either rumble or containing pure tone components, while sound quality analysis may reveal additional information on the market acceptance and the noise influence to customers. The prevailing industrial test standards for product rating of these products, however, have not yet included the concept of sound quality. To investigate the possibility and appropriateness of adopting sound quality analysis in the test standard to rate product noise, we used many personal-computers in accordance with prevailing test standards such as ISO 7779. The results were then analyzed and indexes of sensory pleasantness were calculated by sound quality parameters of loudness, sharpness, roughness and tonality as suggested by Zwicker and Fastl. Comparison between sensory pleasantness values and sound power level indicates a large discrepancy in terms of product rating. It is suggested that the rating of these products by sensory pleasantness may be considered in the future.

4:35

1pNSc7. Noise radiation from a riding mower. Christian M. Skinner (Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804, skinner@sabine.acs.psu.edu) and Courtney B. Burroughs (Penn State Univ., State College, PA 16804)

Riding mowers are a significant contributor to community noise. This has prompted many countries to develop standards to reduce the noise emissions from riding mowers. This research will look at the individual noise sources involved with the riding mower; these are the engine, exhaust, cooling fan, the cutting blades, and the mowing deck housing. The relative contributions to the overall noise from each component were determined by isolating each component during outdoor measurements of the sound power levels. To investigate the mechanisms of the noise generated by the rotating mower blades, a test bed was built to hold the mowing deck and electric motor. The motor was isolated from the test bed via isolation mounts to provide a quiet source of power for the mower deck. Pressure sensors have been embedded into the leading edge of the blades and also at the grass exit to measure the disturbances as the blades and deck interact with each other. Results from these measurements are presented and discussed.

4:50

1pNSc8. International space station acoustics. Jerry R. Goodman (ISS Acoust. Lead, NASA-Johnson Space Ctr., 1814 San Sebastian Ln., Houston, TX 77058, jerry.r.goodman@jsc.nasa.gov)

Control of acoustics is important in the International Space Station (ISS) because of the long-term crew exposure to noise and their confinement. ISS requirements cover modules, government furnished equipment, and payloads. The current status of compliance to these requirements will be described, including examples of where acoustics efforts have aided compliance and where difficulties were experienced. Concerns with obsolete requirements, and lack of acoustic experience and rigor to achieve compliance, will be discussed.

MONDAY AFTERNOON, 4 DECEMBER 2000 PACIFIC SALON E, 2:40 TO 5:25 P.M.

Session 1pNSd

NOISE-CON and Noise: Community Noise and Community-Noise Barriers

Louis C. Sutherland, Cochair
27803 Longhill Drive, Rancho Palos Verdes, California 90275-3908

Brigitte Schulte-Fortkamp, Cochair
Department of Physics/Acoustics, Oldenburg University, Oldenburg D-26111, Germany

Contributed Papers

2:40

1pNSd1. Community noise at Del Mar, California, during two four-day helicopter noise studies. Robert W. Young (710 W. 13th Ave., E-6, Escondido, CA 92025-5511, exy2@worldnet.att.net)

Monday, 30 August 1999, was the first day of a four-day Marine Corps demonstration of noise of helicopters flying north and south along the coast at Del Mar, California. The second four-day demonstration was 26–29 October 1999. A Computer Engineering Limited automatic noise monitor operated continuously at 421 Ocean View Ave. or at 12839 Via Grimaldi. Helicopters flew 0.5- to 5-mile intervals offshore, at altitudes 500 or 2000 feet above sea level. On 2 September 1999, at 421 Ocean View Ave., from midnight to 0800, the one-hour average sound level was less than 41 dB. From 0800 until 0951, helicopter sound levels exceeded an 80-dB threshold 15 times. During the hour ending at 1000, helicopter one-hour average sound level was 72 dB. The maximum fast A-weighted sound level MXFA was 94 dB. The A-weighted sound exposure level ASEL was 103 dB. The 421 Ocean View site is 300 feet east of the shoreline and about 200 feet above sea level. Thus when a helicopter flies 0.5 miles offshore and altitude 500 feet above sea level, it is only 300 feet above and 3500 feet distant from the monitor at 421 Ocean View Ave.

2:55

1pNSd2. The analysis of the environmental influence of drilling installation. Wieslaw Wszolek, Jan Macuda, and Tadeusz Wszolek (Univ. of Mining and Metallurgy, 30-059 Krakow, Poland)

Drilling work, due to its character, settlement, and whole-day service, has potential acoustic danger for the environment. The degree of influence depends on many factors, e.g., the type of drilling equipment, applied acoustic screens, localization of drilling and its surroundings, the type of drilled-trough rocks, and applied drilling technology. The main sources of acoustic noise are motors, pumps, and current generators. The distinction can be made between some procedures that are more noisy than the normal drilling process, for example, tripping. Usually the drilling rig is poorly equipped with noise reduction installation. Experiments were carried out on eight types of drilling assemblies. The received results show that different drilling equipment has an individual influence on the environment. The evident proof of that is the azimuthal acoustic characteristic and additionally the noise level dependence from the cycle of drilling work. The level of LAeq in the surroundings of the examined drilling rig has been made. It has been proved that, due to the noise reduction, the selection of optimal settlement of the drilling rig and optional equipment is possible. In this optimalization process, the cycle of drilling work has been considered. This reduction is possible without additional acoustic protection.

3:10

1pNSd3. Noise produced by religious pilgrimages in Monterrey, Mexico. Fernando J. Elizondo-Garza and Jose de J. Villalobos-Luna (Acoust. Lab., FIME-UANL, P.O. Box 28 ''F,'' Cd. Universitaria, San Nicolas, 66450, N.L., Mexico, fjelizon@ccr.dsi.uanl.mx)

In this paper are shown the result of measurements of the noise produced by religious pilgrimage in the context of the traditional ''Virgin of Guadalupe'' Christian celebration in the city of Monterrey, Mexico. The noise implications are discussed and conclusions are presented.

3:25

1pNSd4. Regulation of amplified sound sources. Eric Zwerling (Rutgers Noise Tech. Assistance Ctr., 14 College Farm Rd., New Brunswick, NJ 08901)

Few sources of noise generate more complaints than amplified music, which can present a difficult enforcement profile if the appropriate approach is not employed. Highly amplified vehicular sound systems are a nuisance increasing in both occurrence and intensity, requiring an enforcement standard that is objective, content-neutral and easily employed by an officer in a police cruiser. A ''plainly audible'' standard meets these criteria and has withstood court challenges. The specific language for inclusion into an ordinance is presented. The regulation, enforcement and curtailment of low-frequency nuisances arising from fixed sources such as bars or parties may prove intractable for performance codes setting A-scale permissible sound level limits. One possible approach, octave band regulation, is not cost-effective and technically feasible for most enforcement agencies which desire efficiently rapid response. An effective approach is a C-scale regulatory protocol prohibiting an amplified sound reproduction device from raising total sound levels within the dwelling of a complainant by specific time-indexed levels. There are technical and evidentiary benefits to taking these measurements within the residence, and not at the property line. This new approach has been recently adopted by St. Augustine, FL and Lafayette, LA. The ordinance language is presented.

3:40

1pNSd5. Verification of sound levels inside discotheques and public show sites. Alberto de Leo, Franco Giuliani, Concetta Fabozzi, and Renzo Tommasi (Natl. Agency for Environ. Protection of Italy (ANPA), Via V. Brancati 48, 00144 Rome, Italy)

In 1997 a decree was issued in Italy with the aim of defining the limits of noise sources inside discotheques and public show sites. ANPA (Italian Agency for Environmental Protection) has started research on about 30 sites in order to know the real levels of exposures to noise inside different types of public show sites and, after, carry out a first verification on proposed limits and methods, and also evaluate the possibility of using the automatic control systems stated by the rules. Such campaigns enable us to obtain elements that are very useful for the development of sectoral rules and for the planning of interventions.

3:55

1pNSd6. Creating noise sources for rotorcraft and aircraft noise simulation models. Kenneth J. Plotkin (Wyle Labs., 2001 Jefferson Davis Hwy., Ste. 701, Arlington, VA 22202, kplotkin@arl.wylelabs.com)

Modern computing power has made noise simulation models, i.e., prediction of actual time histories of sound levels, practical. Switzerland uses its FLULA model for airport noise contours. NASA's new RNM rotorcraft model uses simulation to address the complexities of rotorcraft noise. The Wyle/USAF NMSIM fixed-wing model has been used to address significant effects associated with complex terrain. A requirement for this type of model is the three-dimensional noise source properties of the aircraft being analyzed. This information is generally extracted from fly-over noise tests. Because noise is needed on a complete sphere (or at least a hemisphere), parts of the noise sphere are associated with long propagation distances and/or grazing incidence to measurement positions. This paper reviews the practical problems associated with turning fly-over measurements into source models. Examples from experience with RNM and NMSIM are presented.

4:10

1pNSd7. Rotorcraft noise model (RNM) and acoustic repropagation technique (ART2) validation and application using CH-146 (Bell 412) measurement data. Juliet A. Page and J. Micah Downing (Wyle Labs., 2001 Jefferson Davis Hwy., Ste. 701, Arlington, VA 22202)

A series of acoustic measurements were conducted at Moose Jaw Canadian Forces Base, Saskatchewan, Canada in June, 1998 for obtaining detailed noise data on a large array of ground- and crane-based microphones for a CH-146 helicopter. The project, conducted by the North Atlantic Treaty Organization (NATO) Committee on the Challenges of Modern Society was to define test procedures and analysis methods for the standardization of a NATO rotorcraft noise database. Four independent microphone arrays were deployed by the USAF, NASA, RAF, and DERA. Data from the USAF and NASA arrays were used to validate the Rotorcraft Noise Model (RNM). USAF data were processed via an updated acoustic repropagation technique (ART2) to define the rotorcraft broadband sound hemispheres. The source noise was then propagated to the measurement locations using RNM. Agreement with the original USAF measurements demonstrated consistency of the ART/RNM process. Agreement with the independent NASA measurements provided validation of the RNM as a prediction tool.

4:25

1pNSd8. Investigation of noise reflections using test signals. Lloyd Herman, Jeff Giesey, and Jeremy Ghent (Ohio Univ., 141 Stocker Ctr., Athens, OH 45701)

Repeatable test signals for use with a loudspeaker system were generated to investigate noise reflections from traffic noise barrier walls. Four classes of signals were considered: impulses, maximum-length sequences, pseudorandom binary sequences, and linear frequency-modulated sine waves or chirps. Overall, the chirps yielded the most information from initial experiments. The test system was applied to single and parallel barrier configurations to identify the number of significant reflections, to determine the level of the reflections relative to the direct signal and to quantify the contribution of the reflections to the overall sound level at receivers of interest. For receivers located on the community side of parallel barrier configurations the contribution of the reflections was found to increase with receiver distance from the barrier.

4:40

1pNSd9. Field measurements of a reduced-scale noise barrier's insertion loss and absorption coefficients. Todd A. Busch and Ramon E. Nugent (Acentech, Inc., 1429 Thousand Oaks Blvd., Ste. 200, Thousand Oaks, CA 91362)

In situ testing determined the insertion loss (IL) and absorption coefficients of a candidate absorptive barrier material as part of a project to abate railway noise in Anaheim, California. A 13 000-ft barrier is proposed south of the tracks, but residents of Yorba Linda to the north have had concerns that barrier reflections will increase their noise exposure. To address this issue, a 12-ft-high by 48-ft-long demonstration barrier was built in the parking lot of Edison Field, Anaheim, as part of a public open hours, thereby allowing for acoustical measurements. The barrier's insertion loss (IL) was measured in one-third-octave bands assuming 1/2-scale construction. The IL for three (scaled) railway noise sub-sources (rail/wheel interface, locomotive, and train horn) was measured at six (scaled) distances. The highest total A-weighted IL, after corrections for finite barrier and point-source speaker effects, was 22 dB(A) for rail/wheel noise and 18 dB(A) for locomotive noise. Absorption coefficients were measured using a time-delay spectrometry technique, and were found to be much less than the reverberation room results advertised in the manufacturer's literature, though the resulting increase for noise exposure to Yorba Linda residents is still expected to be less than 1 dB(A).

4:55

1pNSd10. A correction to Maekawa's curve for the insertion loss behind noise barriers. Penelope Menounou (Dept. of Mech. Eng., The Univ. of Texas, Austin, Austin, TX 78712-1063, menounou@mail.utexas.edu)

Maekawa's curve is one of the most established methods for predicting the insertion loss (IL) behind barriers. For the simple case of a barrier modeled as a half plane, the IL is given versus a single parameter, the Fresnel number ($N1$). Predictions obtained by Maekawa's curve deviate largely from experimental data and predictions obtained by analytical solutions when the receiver is either close to the barrier or at the boundary separating the illuminated from the shadow zone. It is shown that if a second Fresnel number ($N2$) is appropriately defined, the IL obtained by the rigorous analytical solutions can be expressed versus $N1$ and $N2$ for several types of incident radiation (plane, cylindrical, and spherical). Accordingly, the single curve in Maekawa's chart can be replaced by a family of curves. Each curve corresponds to a different $N2$ and provides the IL vs $N1$. The Kurze–Anderson formula (a mathematical expression of Maekawa's curve) is also modified to describe this set of curves. The new chart gives predictions that agree well with analytical solutions and experimental data for all receiver locations. Moreover, unlike Maekawa's, this chart can be used for line, point, and plane wave sources. [Work supported by ATP.]

5:10

1pNSd11. Random and periodic square wave barriers in noise control. Nesrin Sarigul-Klijn and Dean Karnopp (Professor and Co-Director of Transportation Noise Control Ctr. (TNCC), Dept. of MAE, UC Davis, Davis, CA 95616-5294, nsarigulklijn@ucdavis.edu)

Federal regulations and increased public interest in reducing transportation noise have led to the construction of miles of sound walls. Most of these walls are almost perfectly non-absorptive barriers with the shortest possible diffractive edge; a straight line. A number of alternate strategies have been studied in recent years, including T-top, Y-top and random edged barriers.Prior investigations by other researchers suggest that it might be possible to improve the performace of a barrier without increasing the average height by introducing a jagged profile. This is because the jagged geometry on the edge of a sound wall alters the sound pressure level in the shadow zone by causing the region of the barrier nearest the receiver to admit multiple paths with variable phase. The direct waves from the diffracting edges of the barrier and waves subsequently reflected from the ground plane are superposed at the receiver causing constructive or destructive interference at the receiver. This is not easily amenable to analytical methods. This led to anechoic chamber testing of the various practical jagged edge treatments. In this paper, theoretical and experimental investigation of the effectiveness of more practical random and square wave barriers are detailed. The TNCC Anechoic chamber experimental facility is utilized for one-sixth scaled barrier tests.

MONDAY AFTERNOON, 4 DECEMBER 2000 PACIFIC SALON B, 2:00 TO 3:35 P.M.

Session 1pSA

Structural Acoustics and Vibration: Methods for Control of Vibration and Radiation

Timothy W. Leishman, Chair
Department of Physics and Astronomy, Brigham Young University, Provo, Utah 84602-4673

Chair's Introduction—2:00

Contributed Papers

2:05

1pSA1. Active noise control studies using the Rayleigh–Ritz method. Senthil V. Gopinathan, Vasundara V. Varadan, and Vijay K. Varadan (CEEAM, Penn State Univ., University Park, PA 16802)

The Rayleigh–Ritz method is used to model a smart composite plate and the piezoelectric patches attached to it. Classical plate theory is used to model the composite plate and electroelastic theory is used to model the segmented piezoelectric patches. Eigenfunctions of a clamped–clamped beam are used as the Ritz functions for the panel and the rigid cavity modes are used to model the acoustic cavity. The dynamic equations of motion for the coupled smart panel-cavity system are derived using Hamilton's principle. The forcing term due to the cavity acoustic pressure is determined by using virtual work considerations. Five collocated pairs of actuator/sensor pairs are attached to the plate at a predetermined placement scheme. A (multi-input–multi-output) MIMO controller is designed for the coupled vibroacoustic system. The output feedback controller is then employed to emulate the optimal controller by solving the Riccati equations. The performance of the control system in reducing the vibration and the noise transmitted into the enclosure is studied. It is shown that the Rayleigh–Ritz procedure can be used more efficiently than the finite-element method for cabin noise control problems.

2:20

1pSA2. Physical limits on the performance of active vibration isolation systems. Vyacheslav M. Ryaboy (Newport Corp., 1791 Deere Ave., Irvine, CA 92606)

Two fundamental issues in active vibration isolation are optimal actuators/sensors placement and optimal control algorithms. In this work, a general approach to active vibration isolation is pursued, with the intent of finding the limitations on the isolation performance that would be valid for all possible active force distributions independent of the particular control algorithms employed. The limiting performance is estimated in terms of isolation efficiency, frequency domain, and work performed by active forces on system displacements, considered a measure of the control effort. These estimates are derived for a general multi-DOF system and expressed as exact inequalities in explicit analytical form. The results show that physical limitations do exist on the efficiency of a generic active vibration isolation system. The principal differences in this respect are discussed between free and supported systems, absolute and relative motions, force and kinematic excitations. Limiting performance estimates can help set design goals for isolation systems. Active force distributions emerging from limiting performance analyses assist in finding rational system configurations and optimal controls. Minimum active work, ex-

pressed in terms of passive subsystem parameters, provides a criterion for optimization of a structure as a potential host for an active vibration control system.

2:35

1pSA3. Active control of structurally radiated sound. Jin Seok Hong (Dept. of Precision Mech. Eng., Grad. School, Hanyang Univ., 17 Haengdang-Dong, Sungdong-Ku, Seoul, Korea, 133-791), Heung Seob Kim (Georgia Inst. of Technol., Atlanta, GA 30332-0405), Kihong Shin, and Jae-Eung Oh (Hanyang Univ., Sungdong-Ku, Seoul, 133-791 Korea)

In this paper, active control of radiating sound (using active structural acoustic control) from a vibrating rectangular plate by a steady-state harmonic point force disturbance is experimentally studied. Two piezoceramic actuators are bonded directly to the panel surface to provide control input structure. The sensor consists of two accelerometers mounted on the plate. Estimated radiated sound signals using a vibro-acoustic path transfer function are used as error signals. The vibro-acoustic path transfer function represents a system between accelerometers and microphones. The controller is a 2 by 2 filtered-x LMS algorithm implemented on a TMS320C30 DSP. The results demonstrate good control of sound radiation using the vibro-acoustic path transfer function.

2:50

1pSA4. Noise reduction in a launch vehicle fairing using actively tuned loudspeakers. Jonathan D. Kemp (Dept. of Mech. Eng., Duke Univ., Durham, NC 27708) and Robert L. Clark (Duke Univ., Durham, NC 27708)

Loudspeakers tuned as optimal acoustic absorbers can significantly reduce damaging, low-frequency, reverberant noise in a full-scale launch vehicle fairing. Irregular geometry, changing payloads, and the compliant nature of the fairing hinder effective implementation of a passively tuned loudspeaker. A method of tuning the loudspeaker dynamics in real time is required to meet the application requirements. Through system identification, the dynamics of the enclosure can be identified and used to tune the dynamics of the loudspeaker for passive dissipation of targeted, high-intensity, low-frequency modes that dominate the acoustic response in the fairing. A loudspeaker model with desired dynamics serves as the reference model in a control law designed to tune the dynamics of a nonideal loudspeaker to act as an optimal tuned absorber. Experimental results indicate that a tuned loudspeaker placed in the nose cone of the fairing dissipates significant acoustic energy, up to 12 dB at the targeted modes, and verifies results calculated from the simulation. [Work supported by AFOSR F49620-98-1-0383 under the supervision of Dr. Daniel J. Segalman.]

3:05

1pSA5. Tuned passive vibration suppression using linear electromechanical coupling. Michael J. Anderson, Tony J. Anderson, William Zornik, Christopher Hocut (Dept. of Mech. Eng., Univ. of Idaho, Moscow, ID 83843-0902), and Jonathan D. Blotter (Idaho State Univ., Pocatello, ID 83209)

One proposed method to suppress vibrations of a mechanical structure is to couple the energy of vibration to an external electric shunt circuit. Coupling of the mechanical energy to the electric circuit is accomplished with a piezoelectric transducer attached to the structure. Significant amounts of mechanical energy can be dissipated by the shunt, but a large inductor is required. In this presentation, we describe an alternative approach that replaces the shunt with a linear passive electric circuit that is connected to another piezoelectric transducer attached to the structure. Even though the electric circuit does not contain an inductor, it is still possible to dissipate significant amounts of mechanical energy in the connecting circuit. The mechanism of energy dissipation in the electric circuit is similar to, but not identical to, that exploited by a tuned electric shunt. Inductance, however, is provided by mutual coupling of piezoelectric transducers on the structure.

3:20

1pSA6. Passive vibration suppression using nonlinear electromechanical coupling. Tony J. Anderson, Michael J. Anderson, Christopher Hocut, William Zornik (Dept. of Mech. Eng., Univ. of Idaho, Moscow, ID 83843-0902), and Jonathan D. Blotter (Idaho State Univ., Pocatello, ID 83209)

The development of a new class of passive vibration and acoustic suppression systems is presented. The approach is to transfer the energy from one mechanical system to another using reversible piezoelectric transducers connected with a passive electric circuit. Response of the first system is suppressed while exciting the response of the second system. The two mechanical systems can be physically separated structures, or different mechanical modes of a single structure. The passive electric circuit is a network containing diodes and/or transistors switched from the piezoelectric voltages. Reductions of 25% in the response of the directly excited system have been shown. This approach is an improvement over a passive shunt technique in that the typically heavy inductor in the electrical shunt is replaced with an existing system mechanical impedance. It also has the advantage over active control techniques in that an external power source is not required and there is no possibility to add energy to the system from the controller.

Session 1pSC

Speech Communication: Learning and Cognitive Processing (Poster Session)

Richard S. McGowan, Chair
CReSS LLC, 1 Seaborn Lane, Lexington, Massachusetts 02420

Contributed Papers

All posters will be on display from 2:00 p.m. to 5:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 2:00 p.m. to 3:30 p.m. and contributors of even-numbered papers will be at their posters from 3:30 p.m. to 5:00 p.m.

1pSC1. The segmental representation of consonant clusters: Constituents, cohabitants, or unitary clusters. Kathleen M. Measer and James R. Sawusch (Dept. of Psych., Park Hall, SUNY at Buffalo, Buffalo, NY 14260)

What is the form of the segmental representation that underlies auditory word recognition? In spoken word recognition, different segmental representations have been proposed for words. Most models of word recognition assume some form of phonetic representation. In a phonetic representation, consonant clusters such as the /spl/ in ''splash'' are treated as a sequence of separate, discrete elements. However, the precise nature of this representation is unknown. Clusters could be composed of abstract phonemes, context-sensitive allophones, or some other unit. A form-based priming paradigm was used with naming, lexical decision, and A–X name matching tasks to investigate the representation of syllable initial clusters. Primes that differ from the targets in the number and order of consonants (in syllable initial clusters) were presented to listeners. For example, primes such as /skrav/, /skav/, /krav/, /sav/, /kav/, /rav/, and /bav/ (control) were used with the target /skrem/. If clusters are composed from their constituent phonemes, the number of phonemes in common between prime and target should determine the magnitude of priming. The results will be discussed in terms of their implications for the segmental representation of speech. [Work supported by NIDCD Grant R01DC00219 to the Univ. at Buffalo.]

1pSC2. Acoustic and perceptual speaker normalization. Patti Adank (Dept. of Linguist., Univ. of Alberta, AB T6G 2E7, Canada), Roeland van Hout (Nijmegen Univ., 6500 HD Nijmegen, The Netherlands), and Roel Smits (Max Planck Inst., 6500 AH Nijmegen, The Netherlands)

An attempt was made to evaluate the performance of several methods for speaker normalization in both the acoustic and the perceptual domain of speech. This was done by comparing the acoustic distributions applied to the same vowel data and further by comparing the acoustic distributions to the perceptual distributions of the vowel data. The normalization methods included, among others, extrinsic (e.g., z-score transformation) and intrinsic methods (Bark transformation), and formant weighting ($F2$) and correction ($F3-F2$). To obtain the acoustic distributions, the normalization methods were applied to $F0$ and formant data from monophthong vowels in /sVs/ context of male and female speakers of Standard Dutch. The perceptual distributions were obtained through an experiment with phonetically trained listeners, whose task was to judge each vowel's height, place of constriction and amount of rounding/spreading. The acoustic and perceptual distributions were compared using correlational- and cluster-analysis techniques. When describing the results of these tests, the focus will be on the extent to which the variation and overlap are the same in the acoustic and perceptual domains, and which normalization methods show a pattern of overlap and variation most similar in both domains. [Work supported by The Netherlands Organization for Research (NWO).]

1pSC3. Talker voice and similarity affect lexical neighborhoods. Liza K. Zimack, James R. Sawusch, Kathleen M. Measer, Paul A. Luce (Dept. of Psych., Park Hall, SUNY at Buffalo, Buffalo, NY 14260, lkz2@acsu.buffalo.edu), and Rochelle S. Newman (Univ. of Iowa, Iowa City, IA 52242)

Words that are similar to a syllable can influence listeners perception of phonemes in the syllable. Known as the lexical neighborhood effect, this has been shown to be robust, but vary in magnitude across talkers. In previous studies [L. Zimack and J. Sawusch, J. Acoust. Soc. Am. **107**, 2918 (2000)], series of nonword syllables were tested in two voices. In one voice, an effect of neighborhood on phonetic perception was clearly found. In the other voice, this effect was absent. The pattern of results suggested that there may be differences in the composition of neighborhoods across different talkers. New studies have explored acoustic–phonetic factors in voices that may underlie this variation. The influence of neighborhood has been examined across syllables from 10 talkers. Measurements of syllable duration, speaking rate, and formant frequencies have also been made for these 10 talkers. Across the talkers, there is variation in the effect of neighborhood on phoneme perception. These data will be presented along with acoustic measurements of vowel spaces and other qualities that may influence the similarity of phonemes and words to one another and thus alter the influence of neighborhood upon perception. [Work supported by NIDCD Grant R01DC00219 to SUNY at Buffalo.]

1pSC4. Segment realization in relation to a carrier word's uniqueness point. Roger Billerey-Mosier (Dept. of Linguist., Univ. of California, Los Angeles, 405 Hilgard Ave., Los Angeles, CA 90095)

This paper determines whether the acoustic realization of English phones differs depending on their position relative to a carrier word's uniqueness point, defined as the earliest point at which that word differs from all the words in the lexicon that share the same initial string. As words are confidently identified at their uniqueness point [Marslen–Wilson (1987)], phonetic material past that point may play a less crucial disambiguating role and show signs of reduction akin to those reported in Wright (1997), who found that vowels in monosyllabic words with relatively few perceptual neighbors are acoustically reduced. The acoustic duration of segments located at and past a carrier word's uniqueness point is compared (e.g., -erry in blueberry and strawberry, respectively, because of blueberry's competitors bluebird, blueblood, etc.). No consistent pattern

of differences is found. This result suggests that the domain of Lindblom's (1990) hypospeech and hyperspeech model is higher than the word. The implications of this result for DeJong's (1995) characterization of stress in English as word-level hyperarticulation are discussed.

1pSC5. The importance of lexical status in word recognition of familiar words and novel items. Evanthia Diakoumakou and Jose Benki (Univ. of Michigan, Prog. in Linguist., 1076 Frieze Bldg., Ann Arbor, MI 48109-1285)

Talker variability and its effects on speech perception is an issue of central importance in word recognition research, and it is known that familiarity with a speaker's voice improves recognition of novel words [Nygaard *et al.*, Psych. Sci. **5**, 42–46 (1994)]. In an investigation of how talker variability interacts with lexical status in the recognition of both familiar and unfamiliar words, listeners performed a closed-set identification task. A visual prime of four potential responses consisting of words and nonsense words that differed minimally phonologically was presented, followed by the audio target stimulus for identification using a response box. The experiment consisted of three 100-trial blocks. The trials were blocked according to talker variability (first block, single talker; second block, multiple talkers; and third block, single talker). Lexical status of the target stimulus varied within trials. Response times were also collected. Listeners performed better when the target item was a word in both the single and the multiple talker conditions. No significant interaction between talker variability and lexical status was found. The results suggest that for listeners lexical status is far more important than talker variability during word recognition.

1pSC6. Distinctive functional load of pitch-accent in Japanese based on word familiarity. Mafuyu Kitahara (NTT Commun. Sci. Labs., 3-1 Morinosato Wakamiya, Atsugi, 2430198 Japan and Indiana Univ.) and Shigeaki Amano (NTT Commun. Sci. Labs., Atsugi, Japan)

The present paper proposes to measure the ''functional load of pitch-accent'' for Japanese words. It is calculated from (1) the ratio of the number of accentually contrastive homophone pairs to the number of noncontrastive pairs; and (2) the difference between word-familiarity ratings for each pair. The functional load is larger when there are more contrastive homophones to noncontrastive homophones for the target word, and when the difference of word familiarity scores is smaller between each homophone pair. A large scale word-familiarity database is used for the calculation of functional load and other statistical properties of pitch accent. Distribution of accent and opposition types at each word length are investigated with respect to familiarity scores. Results show that oppositions which include an unaccented form dominate in short simplex words, from low to high-familiarity range. This suggests that the role of pitch accent in distinguishing homophones is biased to the presence–absence contrast, and not to the location per se of pitch accent. Preliminary results from a perception experiment suggest that top-down information in the lexicon, such as functional load and distribution of accent and opposition types, interacts with the bottom-up process in lexical access.

1pSC7. Lexical tone in spoken word recognition: A view from Mandarin Chinese. Chao-Yang Lee (Dept. of Cognit. and Linguistic Sci., Box 1978, Brown Univ., Providence, RI 02912)

In Mandarin Chinese, segmentally identical words are distinguished by lexical tones, which are realized as distinct fundamental frequency patterns over a syllable. Since tone is used phonemically in this language, the question arises of whether tone is implicated in lexical processes in the same way as segmental structure. Three experiments evaluated the role of Mandarin tones in lexical activation and competition. The form and mediated priming experiments showed that Mandarin listeners exploited lexical tone on-line to disambiguate monosyllabic words that were segmentally identical but tonally distinct. Furthermore, acoustic similarity between tones modulated the magnitude of priming and generated opposite priming patterns between the two priming experiments. The interplay of lexical activation and inhibition, based on a neural network model of lexical access, was invoked to account for the dissociation. The gating experiment showed that more acoustic input was needed to recognize words with tonal minimal pairs, indicative of lexical competition. In conclusion, this study showed that lexical tone in Mandarin Chinese participates in lexical processes, and that the mapping of tonal information onto the lexicon is based on acoustic similarity between tones. The author thanks Sheila Blumstein, Philip Lieberman, James Morgan, Peter Eimas, and Molly Homer for their helpful comments.

1pSC8. Development of prevocalic, intervocalic, and postvocalic /r/ in young American children. Richard S. McGowan (CReSS LLC, 1 Seaborn Pl., Lexington, MA 02420), Susan Nittrouer, and Carol Manning (Boys Town Natl. Res. Hospital, Omaha, NE 68131)

The development of children's /r/ production is the subject of a current investigation of a longitudinal database of children's speech. Nine children were recorded interacting with a parent in approximately 2-month intervals from the age of about 15 months to the age of about 31 months. It would be expected that many of these children would not produce /r/ in an adult manner [e.g., M. M. Vihman and M. Greenlee, J. Speech Hear. Res. **30** (1987)], but the properties of children's production of /r/ as a function syllable position and the production of r-colored vowels has not been examined closely. We have examined two children from this database for their ability to produce /r/ and r-colored vowels in various syllable positions. These children more readily produce postvocalic /r/ and r-colored vowels than they produce prevocalic /r/. Further, it may be that practice with intervocalic /r/ may help in future productions of prevocalic /r/'s. These observations will be quantified using formant frequency and amplitude measures and compared with the productions of /r/ by other children in the database.

1pSC9. Coarticulatory information in natural speech stimuli is crucial for infant recognition of syllable sequences. Suzanne Curtin (Dept. of Linguist., USC, 3601 Watt Way, GFS 301, Los Angeles, CA 90089-1693), Toben Mintz, and Dani Byrd (USC, Los Angeles, CA 90089-1693)

Recent research has determined that coarticulatory information in speech provides important cues to early word segmentation. This experiment investigates whether 7-month-old infants' ability to recognize a string requires the presence of appropriate coarticulatory information in the speech familiarization stream. Following familiarization to a string of CV syllables, infants were tested to determine if sequences that co-occurred in the familiarization string were preferred over those in which the syllables did not appear adjacently during familiarization. Further, the test phase was conducted so that the items had either appropriate or inappropriate coarticulation information. The results indicate that infants tested on items with appropriate coarticulation listened significantly longer to strings that had appeared during familiarization than to the appropriately coarticulated control strings that never occurred together during familiarization. Interestingly, when presented with inappropriate coarticulation test items, infants showed no preference for previously familiarized strings over the non-co-occurring syllable strings. We conclude that infants are sensitive to coarticulation in recognizing sequences in a speech stream. Furthermore, coarticulatory cues, in combination with other cues to segmentation, greatly enhance recognition of syllable sequences. These results suggest that coarticulation plays an important role in early word segmentation. [Work supported by NIH Grants DC-03172 and SSHRC 752-98-0283.]

1pSC10. Infant sensitivity to lexical neighborhoods during word learning. George J. Hollich III, Peter W. Jusczyk (Dept. of Psych., Johns Hopkins Univ., 3400 N. Charles St., Baltimore, MD 21218), and Paul A. Luce (Univ. at Buffalo, Buffalo, NY 14269)

Competition from existing lexical items that share similar phonotactic and phonetic properties could inhibit a child's ability to encode new items. Two studies are reported that examine infant's abilities both to detect the similarity among these ''lexical neighbors,'' words that differ by a single phoneme, and to learn a referent for a novel neighbor after an exposure to a high number of these similar sounding words. In study 1, 15-month-old infants exhibited a novelty preference for a neighborhood prototype after being familiarized with twelve lists of twelve neighbors. This suggests that infants are capable of detecting neighborhood similarity among words. In study 2, 17-month olds were tested on their ability to learn the referents of two novel prototypes after being exposed to their respective lexical neighbors. In the high-density condition, six lists of twelve neighbors were used. The low-density condition consisted of six lists of three neighbors plus nine filler items. Results indicated that word learning was significantly better in the low density condition, both in overall looking times and in infant reaction times to the targeted word. These results fit well with current models of spoken language recognition that suggest a competitive effect for words arising from dense lexical neighborhoods.

1pSC11. 12-month-olds show evidence of a possible-word constraint. Elizabeth K. Johnson, Peter W. Jusczyk (Dept. of Psych., Johns Hopkins Univ., Baltimore, MD 21218, zab@jhu.edu), Anne Cutler (Max-Planck-Inst. for Psycholinguist., Nijmegen, The Netherlands), and Dennis Norris (MRC Cognition and Brain Sci. Unit, Cambridge, UK)

The possible-word constraint (PWC) limits the number of lexical candidates considered for a given input by stipulating that the input should be parsed into a string of feasible words [Norris *et al.*, Cogn. Sci. **34**, 191–243 (1997)]. Any segmentation resulting in impossible word candidates (i.e., an isolated single consonant) is disfavored. Four experiments using the head-turn preference procedure investigated whether 12-month-olds observe the PWC to aid them in word recognition. In the first two experiments, infants were familiarized with lists of words (e.g, ''rush''), then tested on lists of nonsense items containing these words in ''possible'' or ''impossible'' positions [e.g., ''niprush'' (nip+rush—possible) or ''prush'' (p+rush—impossible)]. In the other experiments, 12-month-olds were similarly familiarized with lists of words, but test items occurred in sentential contexts; this condition more readily taxed online segmentation abilities. In the first two experiments, 12-month-olds listened significantly longer to targets in ''possible'' versus ''impossible'' contexts when targets occurred at the end of nonsense items, but not when they occurred at the beginning. In experiments 3 and 4, infants listened significantly longer to words in the ''possible'' condition regardless of target location. These results suggest that 12-month-olds, like adults, use the PWC during online word recognition.

1pSC12. Speech perception and production as evidence for the role of phonological representation in the development of phonological awareness. Virginia A. Mann (Dept. of Cognit. Sci., Univ. of California, Irvine, 3151 Social Science Plaza, Irvine, CA 92697-5100, vmann@uci.edu) and Judith G. Foy (Loyola Marymount Univ., Los Angeles, CA 90045)

Previous research has shown a clear relationship between phonological awareness and early reading ability. The factors underlying the development of phonological awareness are not well understood; both exposure to literacy and the adequacy of underlying phonological representations are cited as antecedents. To test the hypothesis that phonological awareness associates with strength of phonological representations, this study examined speech perception, production, and phonological awareness in 40 four- to six-year-old children attending private preschools. An extensive questionnaire was administered as a probe to home literacy environment; vocabulary letter knowledge and reading ability were also determined. The data indicate that both literacy environment and the adequacy of phonological representation are factors in phonological awareness. They further suggest that strength of phonological representations may be more closely associated with the awareness of onset-rime than with true phoneme awareness. Children with a less developed sense of rhyme had less mature speech perceptual skills and a less mature pattern of articulation independent of age, vocabulary, and letter knowledge. Children with weaker phoneme awareness also showed less mature speech perception and articulation than children with strong abilities, but this appeared to turn on age, letter knowledge, and vocabulary knowledge.

1pSC13. The role of one's own voice in auditory virtual environments. Christoph Pörschmann (Institut für Kommunikationsakustik, Ruhr-Universität Bochum, 44780 Bochum, Germany)

The sound of our own voices contributes significantly to the way we perceive real and virtual environments and has strong implications on the way we speak. It allows speech to be controlled, which is of particular relevance in acoustically ''difficult'' situations, for example when speaking in a noisy environment. Starting from a model of the different pathways relevant to the perception of one's own voice, the implementation of the model in an auditory virtual environment generator is described. The results of an auditory validation experiment show that the naturalness of one's own voice is significantly increased when the model is applied. Furthermore, this system can serve as a tool for psychoacoustic research regarding the perception of one's own voice. In a second experiment the implications of a plausible presentation of one's own voice on the sense of presence in an auditory virtual environment were studied. For this purpose, the virtual environment generator was enhanced to a multiuser system so that it could be used as an auditory teleconferencing system where two or more participants were perceptually placed in one virtual room. Finally, possible commercial applications for the implemented model are identified and briefly described.

1pSC14. Visual influences on internal structure of phonetic categories. Lawrence Brancazio, Joanne L. Miller, and Matthew A. Paré (Dept. of Psych., 125 NI, Northeastern Univ., Boston, MA 02115, brancazio@neu.edu)

We investigated effects of visual place of articulation on internal category structure for voiceless consonants. Volaitis and Miller [J. Acoust. Soc. Am. **92**, 723–735 (1992)], measuring category goodness for tokens from extended auditory voice-onset-time (VOT) /bi/-/pi/-*/pi/ and /gi/-/ki/-*/ki/ continua, found that the best exemplars of /ki/ fell at longer VOTs than those of /pi/, consistent with the well-known difference between /b/-/p/ and /g/-/k/ voicing boundaries. To explore the basis of this best-exemplar shift, we changed perceived place of articulation of auditory stimuli by pairing them with an incongruent visual token (i.e., the McGurk effect). First, we established that the best-exemplar region for /t/ (like /k/) falls at longer VOTs than that for /p/ for auditory bilabial and alveolar continua. Second, subjects rated auditory /bi/-/pi/-*/pi/ stimuli with visual /pi/ (congruent) and /ti/ (incongruent) tokens as exemplars of /pi/ or /ti/, respectively. The best-exemplar region fell at longer VOTs for /ti/ than /pi/, consistent with previously reported visually induced voicing boundary shifts [Green and Kuhl, Percept. Psychophys. **45**, 34–42 (1989); Brancazio *et al.*, J. Acoust. Soc. Am. **106**, 2270 (1999)]. Thus, perceived place of articulation influences internal category structure even when auditory properties are unchanged, suggesting that the structure is not strictly governed by psychoacoustic factors. [Work supported by NIH/NIDCD.]

1pSC15. Auditory-visual context effects on the perception of voicing. Linda W. Norrix (Dept. of Cognit. Sci., Univ. of Arizona, Tucson, AZ 85721) and Connie Keintz (Univ. of Arizona, Tucson, AZ 85721)

Perception of an auditory continuum can be changed by a preceding visual context [K. Green and L. Norrix, in press, JEP: HPP]. This study examines if a visual context for /l/, as in /abla/, influences the voicing distinction for an auditory /aba/-/apa/ continuum. Voice onset time (VOT) and $F1$ onset frequency were measured during production of /aba/, /apa/ (singleton conditions) and /abla/, /apla/ (cluster conditions). Findings indicate longer VOTs in /apla/ compared to /apa/ and lower $F1$ onset frequencies for the cluster compared to singleton conditions. An auditory-only (AO) continuum (/aba/–/apa/) was created and presented to listeners for identification. In an auditory-visual (AV) condition each auditory token was paired with a face saying /abla/ and presented for identification. If subjects use knowledge about coarticulation and its effects on VOT during production, then they should make more /b/ responses in the AV compared to AO condition. In contrast, if they recover $F1$ information from the visual articulation, then more /p/ responses should be obtained. Findings indicate more /p/ responses in the AV compared to AO condition, suggesting that perceivers are recovering information from visual articulations (perhaps tongue height) rather than using stored knowledge about coarticulatory effects on VOT. [Work supported by NSF.]

1pSC16. Identity priming using McGurk stimuli as primes. Iris Oved, Linda W. Norrix, and Merrill F. Garrett (Cognit. Sci. Prog., Univ. of Arizona, Tucson, AZ 85721)

Research indicates that auditory and visual information is integrated during the perception of speech. Conflicting auditory and visual stimuli can result in an illusory experience known as the McGurk effect (e.g., auditory /bav/ dubbed onto a face saying /gav/ results in a perception of ''dav''). This study used a priming paradigm to investigate whether a phonemic representation for the auditory portion of a McGurk stimulus is active *after* the illusory phoneme is experienced. Subjects were given (nonword) prime-target conditions, including: (1) McGurk (e.g., Prime auditory /bav/ + visual /gav/ = ''dav;'' Target auditory /bav/); (2) Incongruent (e.g., Prime auditory-visual /mav/, Target auditory /bav/); (3) Identity (e.g., Prime auditory-visual /yav/, Target auditory /yav/). Results show that mean reaction times to repeat targets were fastest in the identity condition. Response times for the McGurk and incongruent conditions were indistinguishable from one another and significantly slower than the identity condition. This finding suggests that once the auditory and visual information is combined and a phonemic representation is made, the actual auditory signal is no longer available to affect processing of the target. [This work is based on ideas developed by the late Kerry P. Green and supported by NSF.]

1pSC17. Time-varying information for vowel identification in visual speech perception. Deborah A. Yakel (Orange Coast College, 2701 Fairview Rd., Costa Mesa, California, dyakel@occ.cccd.edu) and Lawrence D. Rosenblum (Univ. of CA, Riverside, Riverside, CA 92521)

The relative salience of time-independent and time-varying dimensions during vowel identification accuracy was examined for visual speech perception. Six different conditions were created using nine vowels /o,I,i,e,ae,a,u,U,uh/ in the consonantal context /bVb/: control syllables, variable center syllables, silent center syllables, initial syllables, and final syllables. Control syllables were full syllables. Variable center syllables were produced by deleting the beginning and end transitions, leaving 50%–60% of the vowel. In silent center syllables, 50%–60% of the vowel was removed, leaving the beginning and end transitions. Initial syllables contained the beginning transition, and final syllables contained the end transition. Results revealed that silent center vowels were identified as accurately as the control syllables. All other conditions showed a significant decrease in visual vowel identification accuracy. A follow-up experiment revealed that the silent center syllable accuracy was greater than the variable center syllables even when their durations were similar. These results are analogous to findings in auditory speech perception [J. J. Jenkins, W. Strange, and T. R. Edman, Percept. Psychophys. **33**, 441–450 (1983)] and suggest that the time-varying aspects of the syllable may be the primary source of information for visual vowel identification.

1pSC18. Is what you hear what you see, even in a foreign language? Seung-Jae Moon (Ajou Univ., Suwon, Korea)

Hearing someone's voice, we immediately recognize the owner of the voice if we know that person. Even when we do not know him, we make up a certain mental image of the person. How accurate is the image triggered by the voice alone? Is the relationship between voice and the image triggered by the voice language-specific or universal? These are the issues this study addresses. Korean speech samples from 8 males and 8 females were recorded. Two pictures were taken for each speaker: one showing the whole body with the reference background so that physical characteristics may be easily compared, and the other showing only the face close-ups. 151 non-Asian subjects without any knowledge of Korean language were asked to match the voices with the corresponding pictures. The results were compared with the previous results obtained from 361 native speakers of Korean. While Korean subjects showed highly accurate matching patterns, non-Asian subjects showed much lower accuracy. The tendency to match a certain voice with a certain picture, regardless of being correct or not, was very high with Korean subjects, but much lower with non-Asian subjects. These results seem to suggest that the perception of voice is language-, or culture-specific.

Session 1pSP

Signal Processing in Acoustics: Signal Processing Techniques

David I. Havelock, Chair
National Research Council, IMS/ASP, M36 Montreal Road, Ottawa, Ontario K1A 0R6, Canada

Chair's Introduction—1:25

Contributed Papers

1:30

1pSP1. The role of the time-reversal processor in acoustical signal processing. James Candy and David Chambers (Univ. of California, Lawrence Livermore Natl. Lab., P.O. Box 808, Livermore, CA 94551)

Time-reversal (T/R) processing of noisy ultrasonic sensor array measurements is an established technique for focusing energy in both homogeneous and inhomogeneous media. It can effectively be used to detect flaws or scatterers by utilizing its primary attribute, the ability to focus on the strongest scatterer. In this paper, temporal techniques are discussed that exploit this focusing property to construct signal processing algorithms. First, the underlying theory is briefly discussed illustrating the focusing property and then the discussion of a technique to detect when the T/R processor has focused on the dominant scatterer or flaw follows. Once focus detection is accomplished, that particular scatterer is removed from the measured data enabling a recursive detection of the next scatterer until all have been detected and removed. Next the localization problem is attacked using the dominant scatterer focusing property of the T/R processor. Here a simple nonparametric model of the wave-front is matched with the measured by varying the assumed flaw (scatterer) location until the best match occurs.

1:45

1pSP2. Multiple eigenvalues of the time reversal operator for a small spherical scatterer. David Chambers (Lawrence Livermore Natl. Lab, P.O. Box 808 L-154, Livermore, CA 94551) and Arthur Gautesen (Iowa State Univ., Ames, IA 50011)

Time reverse mirrors have been used to discriminate and focus energy on individual scatterers embedded in difficult media. A time reverse mirror consists of an array of transceivers, each of which records the incident pressure field then re-emits a time reversed image of it. Prada [J. Acoust. Soc. Am. **97**(1) (1995)] described this operation in terms of a time reverse operator whose eigenvalues are associated with individual scatterers in the medium. In particular, it was shown that there was a one-to-one correspondence between the eigenvalues and scatterers for the case of well separated scatterers whose density is identical to the medium. In this talk, Prada's approached is generalized to scatterers of arbitrary densities where the scattered wave is no longer spherical. It is shown that the one-to-one correspondence between scatterers and eigenvalues of the time reversal operator is broken. For the specific case of a small spherical scatterer, there are three eigenvalues. Physical interpretations of the three eigenvalues and the implications for applications of the time reverse method will be discussed. [Work performed under the auspices of the Department of Energy by the Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.]

2:00

1pSP3. A talker tracking strategy for teleconferencing. David I. Havelock (Inst. for Microstructural Sci., Natl. Res. Council, Ottawa, ON K1A 0R6, Canada, david.havelock@nrc.ca)

Sound pickup during teleconferencing is adversely affected by room reverberation. A microphone array steered to the talker can significantly improve the quality of sound pickup but, to follow a conversation in the room, the array must be rapidly steered to alternate talkers. Furthermore, inaccurate and spurious estimates of the source direction must be ''smoothed'' to achieve accurate and stable steering. These two requirements tend to conflict in terms of system response time. One method of achieving both of these performance requirements is to maintain multiple talker trackers simultaneously and to switch between these trackers based on the dynamics of speech and conversation. The array steering performance of a talker tracking system is discussed and results for both simulated and real data are presented.

2:15–2:30 Break

2:30

1pSP4. Analysis filterbank with arbitrary frequency and bandwidth using a multitaper technique. Ryuji Suzuki, John R. Buck (Dept. of ECE & CMaST, UMass Dartmouth, 285 Old Westport Rd., N. Dartmouth, MA 02747), and Antonio H. Costa (UMass Dartmouth, N. Dartmouth, MA 02747)

Acoustical signals are often analyzed for power spectral density. The computational efficiency of the fast Fourier transform (FFT) for the short-time Fourier transform (STFT) is attractive in many contexts, but some applications require nonuniformly spaced analysis frequencies with different bandwidths. Two common examples are the mel frequency scale for speech signals and the constant Q or logarithmic frequency scale analysis of musical signals. These analyses are usually implemented with polyphase filterbanks, wavelets, or an *ad hoc* technique. This research presents a new method that allows arbitrary analysis frequencies and variable bandwidths. The method is based on the multitaper approach from nonparametric spectral estimation. Consequently, this method preserves the benefits of the multitaper approach, namely improved bias and variance properties when compared with the regular STFT-based approach. The multitaper method requires $O(NM)$ computations, where N is the data segment length and M is the number of analysis frequency channels. This is the same order of growth exhibited by the straight discrete Fourier transform with the same N and M. [Work supported by NSF Ocean Sciences.]

2:45

1pSP5. Experimental evaluation of leaky LMS algorithms for active noise reduction in communication headsets. David A. Cartes, Laura R. Ray, and Robert D. Collier (Thayer School of Eng., Dartmouth College, 8000 Cummings Hall, Hanover, NH 03755, david.a.cartes@dartmouth.edu)

An adaptive leaky LMS algorithm has been developed to optimize stability and performance of active noise cancellation systems. The research addresses performance issues related to insufficient excitation, nonstationary noise fields, and signal-to-noise ratio. The algorithm is based on a Lyapunov tuning approach in which three candidate algorithms, each of which is a function of the instantaneous measured reference input, measurement noise variance, and filter length, provide varying degrees of trade-off between stability and performance. Each algorithm is evaluated experimentally for reduction of low-frequency noise in communication

headsets and compared with that of traditional LMS algorithms. Acoustic measurements are made in a specially designed acoustic test cell which is based on the original work of Shaw, Brammer and co-workers and which provides a highly controlled and uniform acoustic environment. The stability and performance of the ANR system, including prototype communication headsets, are investigated for a variety of noise sources ranging from stationary white noise to highly nonstationary measured F-16 aircraft noise over a 20-dB dynamic range. Results demonstrate significant improvements in stability of Lyapunov-tuned LMS algorithms over traditional leaky or nonleaky normalized algorithms while providing noise reduction performance equivalent to that of the NLMS algorithm for idealized noise fields.

3:00

1pSP6. Real-time noise canceling based on spectral minimum detection and diffusive gain factors. Hyoung-Gook Kim, Klaus Obermayer (Dept. of Computer Sci., Tech. Univ. of Berlin, Franklinstr. 28/29, 10857 Berlin, Germany), Mathias Bode, and Dietmar Ruwisch (Cortologic AG, Berlin, Germany)

In this paper, we propose a very simple but highly effective psychoacoustically motivated real-time approach on the basis of spectral minimum detection and diffusive gain factors without a speech activity detector. The first processing step is the calculation of the short-time power spectrum of the noisy speech signal. Estimating the background noise, the system calculates diffusive gain values in real time being obtained in a two-layer structure: Each node of a layer is responsible for a single mode of the power spectrum. The first layer, called the ''minimum detection layer,'' holds the present noise level derived from the minimum of the input power spectrum which is detected within frames smaller than the FFT window. The minimum is transformed into a gain factor function using a signal-to-noise ratio control parameter. The diffusive gain factor interaction of neighboring modes is performed in the second layer, called the ''diffusion layer,'' in order to avoid ''musical tones.'' In the frequency domain, a filtering operation is performed by multiplying the noisy speech power spectrum by the diffusive gain factors to yield the filtered signal spectrum. This latter is transformed to the time domain by an inverse Fourier transform with original noisy phase.

3:15

1pSP7. An improved broadband active noise compressor. Hui Lan, Ming Zhang, and Wee Ser (Ctr. for Signal Processing, School of EEE, Nanyang Technolog. Univ., Singapore 539798)

As the extension of active noise equalization (ANE), an active noise compressor (ANCP) has been proposed recently in the literature to nonlinearly compress the noise power to a certain desired range instead of shifting it linearly. The algorithm uses a variable gain factor to control the dynamic range of the residual noise power. However, the controllability provided by that gain factor is inadequate. In this paper, we propose a new way to represent the gain factor. The new method takes into account the desire to accurately maintain the noise power when it is tolerable and suppress it while exceeding the unbearable level. Simulation and implementation results show that the new method outperforms ANCP in controllability and additionally in computational efficiency [J. W. Feng and W. S. Gan, IEEE Signal Process. Lett. **5**, 11–14 (1998)].

MONDAY AFTERNOON, 4 DECEMBER 2000 PACIFIC SALON C, 1:25 TO 4:45 P.M.

Session 1pUW

Underwater Acoustics: Scattering

Dalcio K. Dacol, Chair
Naval Research Laboratory, Washington, DC 20375-5320

Chair's Introduction—1:25

Contributed Papers

1:30

1pUW1. Standard-target calibration of broadband sonars. Kenneth G. Foote (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

The object of a standard-target calibration is specification of the overall frequency response function of the system. In terms of the transmit and receive signal spectra S_T and S_R, standard-target far-field form function F, and two-way acoustic path loss P for an on-axis far-field calibration; the frequency response function is $H=S_R/(S_TFP)$, where each of the displayed functions is frequency-dependent. In practice, a broadband calibration may be hindered by (1) a vanishing denominator value, causing a divergence in H, or (2) a small numerator value or large denominator value, effectively amplifying the influence of noise. The general solution to these problems is to use multiple dissimilar targets to be able to avoid extrema in F, change the transmit signal waveform to avoid nulls in S_T, and ensure that the signal-to-noise ratio is high over the frequency band of interest. The near-field case can be addressed by adjusting the target range. Given these measures, the standard-target technique demonstrated by Dragonette *et al.* [J. Acoust. Soc. Am. **69**, 1186–1189 (1981)] for scattering measurements may be rendered into general practice for active broadband sonars.

1:45

1pUW2. Bistatic target detection, interferometry, and imaging with an autonomous underwater vehicle platform. Joseph R. Edwards, Henrik Schmidt (Dept. of Ocean Eng., Rm. 5-204, MIT, 77 Massachusetts Ave., Cambridge, MA 02139, jre@mit.edu), and Kevin D. LePage (SACLANT Undersea Res. Ctr., 19138 San Bartolomeo (SP), Italy)

The acoustic detection and classification of completely and partially buried objects in the multipath environment of the coastal ocean presents a major challenge to the underwater acoustics community. However, the rapidly emerging autonomous underwater vehicle (AUV) technology provides the opportunity of exploring entirely new sonar concepts based on mono-, bi-, or multistatic configurations. For example, the medium frequency regime (1–10 kHz) with its bottom penetration advantage may be explored using large synthetic apertures, where acoustic information is accumulated over a series of sonar pings. The performance of such approaches is highly dependent on accurate platform navigation and timing, which poses a significant challenge to AUV developers, particularly because the navigation procedures are themselves dependent on the complicated multipath acoustic environment. Using experimental data from the GOATS'98 SACLANTCEN/MIT experiment, this paper describes an investigation into the feasibility of combining seabed scattering data from

consecutive pings of a fixed parametric source to form a bistatic synthetic aperture for target localization and imaging with an AUV-based receiving platform. The paper describes different levels of bistatic processing including both incoherent and coherent beamforming and very large aperture interferometric approaches, and the associated performance trade-offs are discussed. [Work supported by ONR and SACLANT.]

2:00

1pUW3. Experimental validation of numerical models of 3-D target scattering and reverberation in very shallow water. Irena Veljkovic and Henrik Schmidt (MIT, Cambridge, MA 02139)

OASES 3-D, a wave theory model of seabed insonification, three-dimensional target scattering, and rough seabed reverberation, has been developed to investigate characteristics of multistatic scattering and reverberation from rippled shallow-water seabeds with buried targets. The model combines a scattering theory based on the method of small perturbations with a seismoacoustic propagation model, allowing for field simulations in arbitrarily stratified ocean environments. The validation of this new modeling framework was a major objective of the GOATS'98 experiment carried out in 12–15-m deep water at Elba in May 1998. Spherical and cylindrical targets were buried at different depths within a 10 $\times$10 m area of sandy bottom and insonified by a parametric source. The unique dataset provided by the experiment has been used for both experimental validation of numerical models and identification of features of the 3-D acoustics which distinguish targets from reverberation. The methodology used for estimating the bistatic scattering strength of the targets and the background seabed reverberation from data collected by a moving AUV platform will be described. The performance of the model for predicting scattering from proud and buried targets will be discussed and illustrated by examples including both isotropic and aspect-dependent targets. [Work supported by ONR and SACLANTCEN.]

2:15

1pUW4. The Fermi pseudo-potential and acoustical scattering. Dalcio K. Dacol and Dilip G. Roy (Naval Res. Lab., Washington, DC 20375-5320)

The Fermi pseudo-potential was introduced in quantum mechanics as a means of simplifying problems involving scattering by a multicentered potential including applications to many-body problems. As such, it has direct applications in acoustics. It also provides a physically intuitive basis for discussing scattering by complex objects. In this presentation the pseudo-potential concept is briefly reviewed. Then, applications to scattering by complex objects and to acoustic scattering by multiple objects are discussed. The case of multiple object scattering in an oceanic waveguide will also be discussed in detail. [Work supported by ONR.]

2:30

1pUW5. Extinction theorem for object scattering in a stratified medium. Purnima Ratilal and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

A simple relation for the rate at which energy is extinguished from the incident wave of a far-field point source by an obstacle of arbitrary size and shape in a stratified medium is derived from wave theory. This relation generalizes the classical extinction theorem, or optical theorem, that was originally derived for plane wave scattering in free space and greatly facilitates extinction calculations by eliminating the need to integrate energy flux about the obstacle. The total extinction is shown to be a linear sum of the extinction of each waveguide mode. Each modal extinction involves a sum over all incident modes that are scattered into the extinguished mode and is expressed in terms of the object's plane wave scatter function in the forward azimuth and equivalent plane wave amplitudes of the modes. The only assumptions are that multiple scattering between the object and waveguide boundaries is negligible, and the object lies within a constant sound-speed layer. Calculations for a shallow-water waveguide show that the extinction cross section is highly dependent on measurement geometry, and medium stratification, as well as the scattering properties of the object and may be significantly modified by the presence of absorption in the medium.

2:45

1pUW6. A spectral formulation for the Doppler-shifted field scattered by an object moving in a stratified medium. Yisan Lai and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

A spectral formulation for the 3-D field scattered by an object moving in a stratified medium is derived using full-field wave theory. The derivation is based on Green's theorem for the time-domain scalar wave equation and accounts for Doppler effects induced by target motion as well as source and receiver motion. The formulation is valid when multiple scattering between the object and waveguide boundaries can be neglected and the scattered field can be expressed as a linear function of the object's plane wave scattering function. A normal mode formulation that is more computationally efficient but less general is also derived from first principles. The advantage of the spectral representation is that it incorporates the entire wave number spectrum, including evanescent waves, and so can potentially be used at much closer ranges to the target than the modal formulation. The Doppler effects are illustrated through a number of canonical examples.

3:00–3:15 Break

3:15

1pUW7. On scattering effects due to the proximity and relative position of two bubbles in a sound field. George Kapodistrias and Peter H. Dahl (Appl. Phys. Lab. and Dept. of Mech. Eng., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, georgek@apl.washington.edu)

Recently the authors published results from a theoretical and experimental investigation on scattering of sound from two bubbles symmetrically arranged about the combined beam axis of a set of transducers [J. Acoust. Soc. Am. **107**, 3006–3017 (2000)]. In this presentation the investigation is extended to two additional geometries, with the two bubble array placed at angles of 0° and 45° from the combined beam axis. For each angle, the half interbubble distance d is varied such that the dimensionless variable kd ranges from 0.2–5 (for the 0° case) and 0.2–10 (for the 45° case), where k is the acoustic wave number. Modeling is accomplished by using a closed-form solution derived from the multiple scattering series, with the bubble scattering function expressed in terms of spherical harmonics. Experimental data are obtained by symmetrically arranging two bubbles, each of radius $a \approx 425$ μm, on a fine nylon thread, with the bubbles insonified by tone bursts with a center frequency of 120 kHz. The data closely agree with the simulations, and it is verified that, regardless of the geometry, for $kd \leqslant 1$ the response of the two bubble array drastically departs from the one due to single scattering. This departure is attributed to multiple scattering and is manifested as a reduction in backscattered radiation.

3:30

1pUW8. Rough sea surface limitations on high-frequency SAS imaging. Enson Chang, Ralph E. Chatham, David S. Marx, Matthew A. Nelson, Angela Putney, and L. Kieffer Warman (Dynam. Technology, Inc., 21311 Hawthorne Blvd., Ste. 300, Torrance, CA 90503)

It is a long-held notion that medium-induced signal fluctuations render long-range synthetic aperture sonar (SAS) imaging unfeasible. Recent experimental results in shallow water, however, indicate that near-theoretical SAS performance can be achieved with the aid of (sonar data driven) motion compensation and autofocus algorithms, even for surface- and bottom-reflected paths. This study examined in detail the role of these adaptive compensation algorithms in overcoming rough surface scattering-induced phase errors. Imaging and compensation algorithms were applied

to sonar data that was numerically modeled for moderate to severe surface conditions. We report here the predicted imaging performance as a function of surface conditions and grazing angles. [Work supported by the Office of Naval Research.]

3:45

1pUW9. Modal conversion by rough surface scattering: The key to the *T*-phase. Robert I. Odom, Minkyu Park, and Darin J. Soukup (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

The amplitudes of the propagating acoustic modes associated with *T*-waves decay exponentially below their ray equivalent turning points, and cannot be excited directly by an earthquake. The modal decomposition for a *T*-wave producing earthquake that occurred near the western tip of the Blanco TFZ has been computed. The directly excited higher order modes are characterized by relatively large amplitudes in the ocean crust, significant water-borne components, and often strong interface components at the ocean-bottom boundary. Employing the modal scattering theory of Park and Odom (1999), it is found that energy has been transferred from higher order modes to the Stoneley fundamental and the lower order modes have significant amplitude at the water-bottom interface. Scattering from irregular ocean bottom bathymetry is the mechanism for the energy transfer. The lowest order acoustic modes, modes 1 and 2, are only very weakly excited because they have very small amplitudes at the bottom. This is consistent with the interpretation of de Groot-Hedlin and Orcutt (1999). [Work supported by the NOPP Program and ONR.]

4:00

1pUW10. Integration of numerical scattering functions with the ocean wave-guide. Chee K. Lim and J. T. Goh (DSO Natl. Labs., 20 Science Park Dr., Singapore Science Park, Singapore 118230, lcheekho@dso.org.sg)

Based on the generalized modal formulation of acoustic scattering by F. Ingenito [J. Acoust. Soc. Am. **82**, 2051–2059 (1987)] in a stratified ocean, we present a method for integrating a numerical scattering function obtained from a commercial boundary element code, SYSNOISE, with an acoustic propagation code to study scattering from arbitrary objects. We benchmarked the hybrid method by comparing the scattered fields for a rigid spherical shell obtained using this approach with an analytical solution. We discuss issues related to the numerical implementation of the approach, such as the extraction of the scattering function for an unbounded ocean medium, correct extraction of multiple modal scattering kernal using a single scattering function (for axis-symmetrical object) and numerical phase compensation procedures associated with numerical source location. The methodology developed here would allow the study of acoustic scattering problems in the ocean for arbitrary-shaped elastic objects, which are not amenable to analytical means but solvable only by numerical methods.

4:15

1pUW11. Comparison of the shallow water acoustic scattering from a submarine-shaped object with that from canonical shapes. Chee K. Lim (DSO Natl. Labs., 20 Science Park Dr., Singapore Science Park, Singapore 118230, lcheekho@dso.org.sg)

The scattering from an arbitrary-shaped object like a submarine is a complex process. In practice, for detection studies, canonical shapes such as a cylinder have been widely used to emulate a submarine-shaped object. However, the validity of this approximation is not easily quantified. In a previous paper [Chee K. Lim, ''Integration of numerical scattering functions with the ocean wave-guide,'' 139th ASA, Nov. 2000], the approach to integrating scattering functions computed using commercial codes with a propagation code was presented. The approach allows convenient calculation of scattering from very complex elastic objects. This presentation will compare the acoustic scattered field of eight different targets—a rigid sphere, rigid cylinder, rigid cone, rigid submarine-shaped object, elastic cylinder, stiffened elastic cylinder, stiffened elastic cone and, lastly, a stiffened elastic submarine-shaped object. Five aspects source angles (0, 45, 90, 135 and 180 deg) and their corresponding forward- and backward-scattered field will be compared using the results of the submarine-shaped object as the reference solution. The results would provide useful insights for future underwater acoustic scattering modeling and model-based signal processing.

4:30

1pUW12. Scattering of acoustic waves by a prolate body in plane-stratified waveguide. Boris G. Katsnelson, Valery A. Grigoryev (Voronezh Univ., 1 Universitetskaya Sq., Voronezh 394593, Russia), Venedict M. Kuzkin, and Valery G. Petnikov (General Phys. Inst., Moscow 117942, Russia)

Technique for calculation of diffraction field by spatially localized inhomogeneity is proposed. This technique is based on the fact that the scattering matrix of the waveguide mode is expressed through the scattering amplitude in free space. This permits separation of the diffraction problem in free space and the propagation problem in the waveguide (the problem of calculation of waveguide modes). Requirements for sound-speed variation and for distance between scatterer and waveguide boundaries are formulated. On the basis of the energetic approach, the concept of the cross section in the waveguide is proposed. The coupling modes and sound absorption are discussed as well. The space-time structure of a sound field scattered by the moving localized inhomogeneity in the direction of incident wave (the forward scattering) is given and discussed. As an example, structure of the sound field for a scatterer modeled by spheroid prolate for different parameters of waveguides is calculated. [Work supported by RFBR, Grant 99-02-17671.]

Additional registration fee required to attend this Tutorial.

Session 1eID

Interdisciplinary: Tutorial Lecture on Virtual Musical Instruments

Joseph Pope, Chair
Pope Engineering Company, P.O. Box 590236, Newton Centre, Massachusetts 02459-0002

Chair's Introduction—7:00

Invited Paper

7:05

1eID1. Virtual musical instruments. Julius O. Smith III (Ctr. for Computer Res. in Music and Acoust., Dept. of Music, Stanford Univ., Stanford, CA 94305)

Virtual musical instruments are interesting from multiple points of view. The listener may, or may not, know that the music is computer generated. The composer can use the computer to produce sonorities that cannot be obtained in a real environment, such as strings many meters long, or instruments that gradually change form. The performer can benefit from virtual instruments because they are usually easier to play than real instruments. The scientists, being able to reproduce the sonority of a particular instrument, show that the physics of the instrument itself has been understood. For these reasons and more, virtual musical instruments based on mathematical acoustics have been successful for many years in the computer music research community. In this tutorial we will cover some of the more effective virtual instrument algorithms, and how they can be played and varied in real time. Examples will be described and sound examples played. Starting from the simplest cases of plucked and struck strings, we will show how simple configurations of delay lines, digital filters, and nonlinear elements can be used to synthesize realistic sounding musical instruments. Examples will work up to bowed strings and other ''self-sustained oscillators'' such as in the flute and clarinet. The richness and breadth of the timbral spaces offered by these models will be illustrated.

Session 2aAA

Architectural Acoustics and Engineering Acoustics: Loudspeaker Student Design Competition

Neil A. Shaw, Cochair
Menlo Scientific Acoustics, Inc., P.O. Box 1610, Topanga, California 90290-1610

Allan Devantier, Cochair
Infinity Engineering, Harman Consumer Group, 8500 Balboa Boulevard, Northridge, California 91329

The ASA Technical Committees on Engineering Acoustics and Architectural Acoustics, in association with the American Loudspeaker Association, JBL Professional, JBL Consumer, Infinity and Revel, are sponsoring a Student Design Competition, which will be professionally judged at this meeting. The purpose of this design competition is to encourage students enrolled in Engineering, Physics, and other University curriculums that involve engineering, physics, materials, electroacoustics and/or acoustics to express their knowledge of electroacoustic transducer design. This knowledge will be demonstrated in the prototype design of an electroacoustic transduction system where various design parameters must be balanced to provide a product that meets the design objectives. The 2000 competition involves the design of a loudspeaker system, one specimen of which is to be developed, that may be placed on a pedestal or surface mounted to a room surface with physical and performance characteristics defined by the Design Specification.

The submitted designs will be judged by a panel of professional loudspeaker designers and design consultants at the meeting. An award of $1,000 US will be made to the submitter(s) of the entry judged ''First Honors.'' Four awards of $500 US each will be made to the submitters of entries judged ''Commendation.''

Entries will be on a display in this session. Each entry consists of a poster presentation and a prototype specimen device. In addition, the results of the objective and subjective tests performed for each entry will also be presented.

Session 2aAO

Acoustical Oceanography: Bioacoustics III

Timothy K. Stanton, Chair
Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Bigelow 20, Woods Hole, Massachusetts 02543-1053

Chair's Introduction—7:50

Contributed Papers

8:00

2aAO1. Acoustic investigation of the Hawaiian mesopelagic boundary community. Kelly Benoit Bird and Whitlow Au (Hawaii Inst. of Marine Biol., 46-700 Lilipuna Rd., Kaneohe, HI 96744)

A 200 kHz echosounder modified to sample directly into a computer was used to survey the mesopelagic boundary community (MBC) in waters off the leeward coasts of three Hawaiian Islands from dusk to dawn. The target strengths of individual mesopelagic organisms captured in Isaacs-Kidd midwater trawls were measured with the echosounder in a shipboard seawater tank for use in echo-energy integration calculations. Acoustic survey data were mapped in ArcView, allowing analysis of spatial structure. The MBC is spatially heterogeneous both horizontally and vertically. The parameters of patch structure vary significantly throughout a night and between islands. The density of mesopelagic organisms was consistently higher off the coasts of Oahu and Lanai than the coast of Hawaii. However, the density of organisms observed in all locations was high, reaching a maximum of 1800 organisms/m^3 off Oahu and 700 organisms/m^3 off Hawaii. Patterns in density were significantly affected by the location's distance from the shoreline, but not by its depth. Patterns in density suggest that the MBC undergoes a significant diel horizontal migration in addition to its well-documented vertical migration. Due to this predictable and conserved horizontal migration, the MBC is found much closer to shore than previously reported.

8:15

2aAO2. Acoustical study of the prey field and supporting food web for whales in the Ligurian Sea north of Corsica. Duncan E. McGehee (BAE Systems, 4669 Murphy Canyon Rd., Ste. 102, San Diego, CA 92123, duncan.mcgehee@baesystems.com) and David A. Demer (Southwest Fisheries Sci. Ctr., La Jolla, CA 92037)

The Sound, Oceanography, and Living Marine Resources (SOLMAR) program is a study coordinated by the NATO SACLANT Undersea Research Centre to develop ways of mitigating risks to whales from anthropogenic sound sources. One objective of the SOLMAR program is to

examine the correlation between the distribution of whales in parts of the Mediterranean Sea and that of their prey, lower trophic levels, and basic oceanographic parameters. A multiship field study was conducted in the Ligurian Sea north of Corsica in August 1999, as will be a follow-up study in August 2000. The primary role of the R/V AMMIRAGLIO MAGNAGHI is to map basic physical oceanographic parameters with a CTD; phytoplankton with a fluorometer; small zoo-plankton (copepods and smaller) with a six-frequency TAPS (Tracor Acoustic Profiling System) and small plankton net; and large zooplankton (euphausiids and larger) with a 120-kHz split-beam echosounder. Measurements in 1999 were consistent with a counterclockwise circulation in the basin causing Ekman pumping of nutrients in the center. Phytoplankton were concentrated in the center, small zooplankton were around the periphery, and larger zooplankton were associated with the shelf break and basin center. Results from the two field seasons will be discussed and compared.

8:30

2aAO3. Biophysical interactions at the Almeria–Oran front. Sophie Fielding, Nick Crisp, and Alex Mustard (Southampton Oceanogr. Ctr., European Way, Southampton SO14 3ZH, UK)

The Almeria–Oran front forms where Atlantic origin waters meet Mediterranean surface waters at the eastern end of the Alboran Sea. Following a multidisciplinary cruise on RRS DISCOVERY in December 1996, analysis of temperature and salinity on density surfaces at the front showed that periodic subduction of Mediterranean surface waters (MSW) influenced the distribution of phytoplankton and zooplankton in the frontal region. A layer of fluorescence coincident with the subducted MSW indicated that phytoplankton were drawn down and along isopycnals to depths of 200 m. The study of scattering layers, identified with acoustic backscatter data from the shipboard ADCP and a SIMRAD EK500 echosounder, revealed that a layer of zooplankton was also found coincident with the drawn-down phytoplankton. This layer persisted during and despite diel vertical migration. Smaller zooplankton, not undertaking diel vertical migration, were concentrated in the fast-flowing frontal jet as seen in high-resolution OPC data. EK500 target-strength data indicate that the subducted communities may constitute different size classes to those in the surrounding waters. Samples from Longhurst Hardy plankton recorder tows across the front provide additional ground truthing for our acoustic observations. [Work supported by EU MAST, NERC, and DERA.]

8:45

2aAO4. Changes in volume reverberation from deep to shallow water in the Gulf of Mexico. Richard H. Love (BayouAcoust., 209 Baywood Dr., Pass Christian, MS 39571), Charles H. Thompson, and Redwood W. Nero (Stennis Space Center, MS 39529)

Volume reverberation at frequencies below 10 kHz is caused primarily by resonance scattering from swimbladders of dispersed fish. Since fish populations change from deep to shallow water, the character of volume reverberation should likewise change. However, there are few data available from shelf and slope waters to document expected changes. Therefore, the Naval Research Laboratory conducted a short experiment to examine volume reverberation in deep, slope, and shelf waters of the eastern Gulf of Mexico. Day and night measurements at frequencies from 1 to 10 kHz were made with shallow explosive sources and downward-looking receivers in water depths of 200, 450, and over 3000 m. Reverberation at the deep location was typical of that caused by small mesopelagic fishes of the ''deep scattering layer;'' scattering strengths increased rapidly with frequency up to about 5 kHz and were relatively constant at higher frequencies. At the shelf location, scattering strengths were low during the day and increased by 20 dB at 10 kHz to 40 dB below 2 kHz, as larger fishes rose off bottom at night. Reverberation over the slope combined characteristics of the deep and shelf locations. [Work supported by ONR.]

9:00

2aAO5. Herring hydroglyphics in littoral waters of the northern Gulf of Mexico. Redwood W. Nero, Charles H. Thompson, John R. Dubberley (Naval Res. Lab., Stennis Space Center, MS 39529-5004, woody.nero@nrlssc.navy.mil), and Richard H. Love (Bayou Acoust., Pass Christian, MS 39571-2111)

A large shoal of fish occurring in the vicinity of the 220-m isobath was observed using a standard 38-kHz fisheries echosounder and a 1.5- to 10-kHz low-frequency fish sonar (LFFS) for several days in July 2000. The fish behaved like herring, exhibiting a rapid rise to the sea surface at dawn, formation into schools, and a rapid descent to the sea floor. Schools remained at depth through the day and gradually rose to the sea surface at dusk and then rapidly descended and spread out into a diffuse scattering layer at 75- to 125-m depth. Shifts in resonance frequencies during migration, release of gas bubbles during migration, and strong avoidance of the vessel when maneuvering, all suggest the fish were most likely round herring, *Etrumeus teres*, which are common at these depths in the NMFS historical trawl survey data. An examination is made of some of the scattering characteristics of the schools and layers of these fish and comparisons of the 38-kHz data to scattering at 500-Hz bands from 1.5 to 5 kHz and 1-kHz bands from 5 to 10 kHz. [Work supported by ONR.]

9:15

2aAO6. Monitoring fish movement using an ADCP. Ranjan Patro, Len Zedel, and Cristina Spanu-Tollefsen (Dept. of Phys. and Physical Oceanogr., Memorial Univ. of New Foundland, St. John's, NF A1B 3X7, Canada)

An Acoustic Doppler Current Profiler, ADCP, can detect the presence of fish in water using the backscatter intensity. The Doppler profiler does not, however, make a point measurement; rather, measurement is made by averaging multiple beams and assuming that the current velocity is uniform between the distinct sample locations. As a result, individual fish speed cannot be measured. However, data are presented that demonstrate that the ADCP can measure the swimming speed of large fish schools. Fish speed and direction were measured for Norwegian spring herring. Observed speeds were 0–40 cm sec^{-1}. Diel vertical migrations were observed with Norwegian herring ascending to the surface at dusk, and descending to greater depth at dawn. The accuracy and precision of an ADCP is a complex function of many variables (i.e., frequency, pulse length, transducer characteristics, backscatter strength, type and distribution of scatters). The other problems to be considered are sampling criteria and calibration. These sampling problems are explored for the case of ADCP measurements of fish movement. [Work supported by the Natural Sciences and Engineering Research Council of Canada and an Atlantic Canada Opportunities Agency Infrastructure grant.]

9:30

2aAO7. Comparison of measured and calculated absorption spectra due to ensembles of fish with swim bladders. Orest Diachok (Naval Res. Lab., Washington, DC 20375)

Absorption spectra derived from long-term, broadband transmission loss measurements were compared with calculated absorption spectra due to ensembles of identical fish with swim bladders over the frequency range 0.5–5.0 kHz, and near-coincident trawling data. The dominant fishes during this experiment were sardines ($\sim$60%) and anchovies ($\sim$30%). Only nighttime data, when the separation between these fishes is large (compared to wavelength), were considered. Measured distributions of the lengths of sardines and anchovies were separated into groups, which were of nearly uniform length—essentially year classes. Matching of measured and calculated absorption spectra yielded estimates of the number densities and Q's of species and year classes. The spatial number densities of adult sardines, which were derived from absorptivity measurements ($1.1/m^2$) and nearly coincident echo sounder data ($0.6-1.6/m^2$), were consistent. The Q's and resonance frequencies of the absorption lines, which were attributed to adult and juvenile sardines, were approximately

2.3 at 1.3 kHz and 4.7 at 3.8 kHz, respectively. These values of Q appear to be dominated by the Q_0 of individual fish. Limited laboratory measurements of Q_0 of other species versus fish age are consistent with the observed age dependence. [Work supported by ONR.]

9:45

2aAO8. Species identification of pelagic fish schools using acoustic descriptors and ancillary information. Gareth Lawson[a] (Marine and Coastal Management, Private Bag X2, 8012 Rogge Bay, South Africa), Manuel Barange (Plymouth Marine Lab., Plymouth PL1 3DH, UK), and Pierre Freon (Institut de Recherche pour le Developpement, France)

It is shown that schools of similarly sized and behaving pelagic fish species can be accurately identified to species over a large spatial area, on the order of the broad spatial-scales typical of standard assessment surveys, and over a time-span of three years. Acoustic measurements of morphometric, energetic, and bathymetric features of anchovy, sardine, and round herring schools were extracted using commercially available software from acoustic data collected by a conventional single-frequency narrow-band echosounder during trawling operations of pelagic stock surveys of the South African continental shelf (November 1997, 1998 and 1999). Discriminant function analysis of 18 descriptors of the 214 schools for which the species composition was satisfactorily determined by trawl samples indicated that schools could be correctly identified to species in 88.3% of all cases (94.9% for anchovy, 82.6% sardine, 82.6% round herring). The three species were differentiated primarily on the basis of school bathymetric position and backscattered energy. Including ancillary information (latitude, longitude, sea surface temperature, bottom depth, and time of day) in analysis improved the accuracy of school identification to 94.9%, since such variables allowed discrimination on the basis of interspecific differences in habitat use, as well as in schooling behavior.
[a] Currently at MIT/Woods Hole Joint Program.

10:00–10:15 Break

10:15

2aAO9. Estimating the scattering contribution from microstructure and marine organisms using multiple-frequency acoustic data collected from a Gulf of Maine internal wave. Joseph D. Warren, Peter H. Wiebe (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), Harvey E. Seim (Univ. of North Carolina, Chapel Hill, NC 27599), and Timothy K. Stanton (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

High-frequency sound (>10 kHz) is scattered in the ocean by many different sources. Marine organisms are often thought to be the primary source of backscattered sound energy. Because of this, acoustical methods are used widely to quickly survey large areas of the ocean for the distribution and abundance of animals. Field experiments and recent theoretical work suggest that temperature and salinity microstructure in some oceanic regions could cause acoustic scattering at levels comparable to that caused by organisms. A method which uses multiple-frequency acoustic data to discriminate between biological and physical sources of scattering has been developed and applied to data collected in an internal wave in the Gulf of Maine. Model calculations show that in certain regions, scattering contributions from turbulence are equal to the contributions from animals. Using net tow information to account for the scattering from biological sources, acoustic scattering data combined with temperature and salinity profiles are inverted for estimates of the dissipation rate of turbulent kinetic energy (ϵ). The estimates of ϵ made at the depths where an internal wave occurred are several orders of magnitude higher than the ϵ estimates from depths above and below the internal wave.

10:30

2aAO10. Comparison and evaluation of methods for distinguishing among several functional classes of zooplankton and fish using acoustic backscatter data at four frequencies. Gordon L. Swartzman (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105), D. Van Holliday (BAE SYSTEMS, San Diego, CA 92123), Kenneth O. Coyle (Univ. of Alaska, Fairbanks, AK), and Jeffrey M. Napp (Alaska Fisheries Sci. Ctr., Seattle, WA 98115)

Zooplankton net samples data were collected at 30 locations near the Pribilof Islands, Alaska in September 1999 using a MOCNESS. Simultaneous volume backscatter data were collected with an HTI echosounder system at: 43, 120, 200, and 420 kHz. Three algorithms were compared for their ability to classify acoustic data into the dominant zooplankton functional/size groups (euphausiids, copepods, and ostra cods) and separating them from fish. The algorithms tested include: (1) combining morphological image processing and image differences to identify patches in different size ranges, and using the forward problem calibrated to plankton found in MOCNESS hauls to ascribe size/functional groups to the different patches; (2) application of canonical correlation between acoustics (backscatter volume and target strengths) and plankton biomass captured by the MOCNESS; and (3) application of inverse techniques to estimate the number of individuals in set size classes using multiple theoretical models of backscatter volume compared to plankton biomass captured by a net. Knowledge of the strengths and weaknesses of each technique allows us to better interpret broad-scale acoustic survey results from the same cruise, thus providing a synoptic view of the fish and their plankton prey which are in the size and density range to be acoustically detected.

10:45

2aAO11. Validity study of the distorted wave Born approximation (DWBA) model: Application to the acoustic scattering by marine organisms. Dezhang Chu and Timothy Stanton (Dept. of Appl. Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Applications of the distorted wave Born approximation (DWBA) to the acoustic scattering by weakly scattering marine organisms have shown much progress [Chu *et al.*, J. Acoust. Soc. Am. **93**, 2985–2988 (1993); Stanton and Chu, *ibid.* **103**, 236–253 (1998)]. Although the applicability of the model has been examined for simple objects such as spheres, cylinders, and prolate spheroids, the validity of applying the DWBA model to the acoustic scattering by marine organisms with more complicated shapes has not been studied. To evaluate the performance of the DWBA model, a systematic study of the DWBA will be presented. An in-depth study of the 1D problem reveals the fundamental difference between the exact and DWBA solutions and helps us better understand the scattering problem. By studying the internal field of 1D, 2D, and 3D problems for a more general bistatic geometry and evaluating the volume integral derived from the exact Helmholtz–Kirchhoff integral equation, we are able to estimate the validity of the DWBA model in describing the acoustic scattering by arbitrarily shaped marine organisms. [Work supported by NSF and ONR.]

11:00

2aAO12. Wavelet analysis comparing spectral characteristics of high frequency acoustics and hydrographic properties on Georges Bank. Karen Fisher (2154 Snee Hall, Cornell Univ., Ithaca, NY 14853, kef10@cornell.edu) and Peter Wiebe (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Acoustic surveys of Georges Bank were carried out during the GLOBEC Broadscale program in January, March, and June of 1998 and 1999. The Greene Bomber towfish collected temperature, salinity, and fluorescence as well as the acoustics data at 120 and 420 kHz, from approximately 3 meters depth on all six cruises. With over 1200 kilometers of acoustic trackline on each cruise, covering well mixed, frontal, and well stratified water columns, a method is needed to characterize patch structures encountered. Wavelet analysis offers a computationally effective approach that can be applied to all the data obtained along the trackline.

Variance analysis of the wavelet transform provides both spatial and spectral resolution simultaneously, allowing distinct consideration of spectra obtained in any spatial subset of the entire dataset. Temporal and spatial variations in the characteristic slopes and length scales of temperature, salinity, and fluorescence indicate that distinct spectral signatures exist that can be exploited to sort out physical and biological contributions. The extent to which physical and biological spectral signatures influence the wavelet variance spectra of the acoustic data is examined in each hydrographic regime.

11:15

2aAO13. Volume acoustic backscattering measurements with a 68-kHz cylindrical array. Timothy C. Gallaudet and Christian P. de (Marine Physical Lab., Scripps Inst. of Oceanogr., UCSD (Code 0205), 9500 Gilman Dr., La Jolla, CA 92037-0205, tcg@mpl.ucsd.edu) Moustier

Most sensors designed to measure volume acoustic backscatter over the entire water column have limited horizontal coverage. The 68-kHz cylindrical arrays of the U.S. Navy's Toroidal Volume Search Sonar (TVSS), towed 80 m below the sea surface in 200-m water depth, have been used to measure volume backscatter over the entire water column with about 4.5 deg of angular resolution over a swath roughly 300 m wide. The data were analyzed to determine the volume acoustic backscattering strengths and cross sections, and the target strengths of (1) near-surface bubble clouds in the tow ship wake; (2) midcolumn zooplankton layers; and (3) near-bottom fish schools. Unlike previous acoustic studies of ship wakes, these results are obtained from data which sample the bubble field at a constant distance from the ship. Echo integration over a midcolumn scattering layer is used to assess the patchiness of zooplankton in the region. Target strengths of near-bottom fish schools are easily distinguished from the surrounding seafloor backscattering strengths. [Work sponsored by ONR-NRL (Contract No. N00014-96-1-G913).]

11:30–12:00

Panel Discussion

TUESDAY MORNING, 5 DECEMBER 2000 PACIFIC SALON D, 7:30 A.M. TO 12:05 P.M.

Session 2aBB

Biomedical Ultrasound/Bioresponse to Vibration, Physical Acoustics and Acoustical Oceanography: Detection and Characterization of Bubbles, Acoustic Cavitation, and Associated Physical Effects I

Ronald A. Roy, Cochair
Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Timothy G. Leighton, Cochair
Institute of Sound and Vibration Research, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom

Chair's Introduction—7:30

Invited Papers

7:35

2aBB1. Methods for detecting bubbles and monitoring cavitation activity; a report from Working Group 22 of Standards Committee S1. Wesley L. Nyborg (Phys. Dept., Univ. of Vermont, Burlington, VT 05405)

Few liquids or solids are homogeneous; most contain cavities filled with air or other gas. They may be large or small, desirable or harmful. In diagnostic medicine, small gas-containing particles are introduced into the circulation as contrast agents to improve ultrasound images. In lakes and oceans, bubbles produced at the surface present difficulties for sound propagation. In industries gaseous phase is sometimes desired in manufactured products, but unwanted gas bubbles can be a serious problem. Exposing cavities to sound produces a complex panoply of activity which includes simple breathing oscillations, shape oscillations, high-speed travel, jet formation, dramatic implosions, fragmentation, bubble growth and microstreaming. Mechanical stresses and chemicals produced by acoustic cavitation are capable of causing a host of physical, chemical and biological effects which can be desired, as in sonic cleaning or sonoporation, or harmful if unwanted chemical action or solid erosion is produced. The ASA Committee on Standards established a working group to prepare a technical report dealing with the many techniques which have been used to characterize bubbles and monitor cavitation activity. In this report, 25 techniques are described. Principles of operation are explained and applications discussed for each of these.

8:00

2aBB2. Measurement of oceanic bubble populations using broadband acoustics. Eric J. Terrill (Scripps Inst. of Oceanogr., Mail Code 0213, La Jolla, CA 92093-0213, et@mpl.ucsd.edu)

The size distribution of bubbles in the ocean will evolve with a number of different temporal and spatial scales as a result of both the physics which govern their formation (breaking wind waves, breaking waves in the surf zone, or ship wakes) and the physical processes which control their lifetimes: turbulent mixing, bubble rise speed, and gas dissolution. A broadband acoustic bubble sizing technique that is suitable for deployment in the rigors of the ocean environment has been developed and deployed on a number of occasions to study the physics of the bubble field and their implications to underwater sound propagation (Terrill and Melville, 2000).

The technique is based upon inversions of direct acoustic attenuation and sound-speed measurements at frequencies ranging from approximately 2–200 kHz using broadband pulse transmissions at ping rates up to 25 Hz across short path lengths of O (0.1–1) m. The accuracy of the measured bubble size distributions has been confirmed with both internal consistency checks based on the complex dispersion relationship as well as laboratory comparisons with independent optical techniques. A review of the development of the bubble sizing instrumentation will be presented and results from a number of different field deployments discussed.

8:25

2aBB3. Bubble characterization in the surf zone. Grant Deane (Scripps Inst. of Oceanogr., UCSD, La Jolla, CA 92093-0238)

Breaking waves entrain bubbles ranging in radius from tens of microns to centimeters. In the surf zone, these bubbles are evident above the ocean's surface as patches of foam that persist for a few seconds after the passage of a breaking wave. Beneath the surface, the bubbles are organized into dense plumes with void fractions of air exceeding 0.1. In the absence of additional breaking, the void fraction of air decreases to 1e-6 or less in a minute or so as bubbles rise to the surface, dissolve, and disperse. Thus, a full characterization of bubbles in the surf zone requires instruments that operate over a wide range of void fractions and bubble radii. In practice, a variety of techniques is used, and the application and results from a number of optical and acoustical instruments, including photographic techniques, travel time sonars, Doppler sonars, and acoustical resonators will be discussed. Some of the physical acoustical effects associated with the large populations of bubbles entrained by breaking surf, such as acoustic hot spots in surf noise and acoustic transmission dropouts, will also be presented.

8:50

2aBB4. On bubble-mediated scattering from the sea surface: Modeling, and field and laboratory measurements. Peter H. Dahl (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

Bubbles, located just below the air–sea interface, contribute significantly to the observed level of acoustic scattering originating from the sea surface, in contrast, for example, to microwave scattering. Typically, such bubbles are the result of breaking wind-generated waves, and thus their concentration is closely linked to wind speed. In this paper we discuss scattering from bubbles located within the proximity of the sea surface, with emphasis on the effects of the nearby reflecting surface, which gives rise to multiple paths from source to bubble to receiver. The effects of scattering geometry, e.g., as in bistatic forward scattering and monostatic back-scattering, are illustrated using field data taken at a frequency of 30 kHz. An interpretive model for the contribution of near-surface bubbles to the apparent scattering cross section per unit area of sea surface is also discussed. This model has both a monostatic and a bistatic form, and, somewhat paradoxically, the bistatic form does not reduce to the monostatic form in the limit of source and receiver co-location. The issue is clarified, however, by examining the constituent pressures associated with the multiple paths, and by well-controlled laboratory measurements that include rough surface effects. These confirm both monostatic and bistatic forms of the model.

9:15

2aBB5. Acoustic cavitation and sonoluminescence. Thomas J. Matula (Appl. Phys. Lab, Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

The violent collapse of a bubble in an acoustic field can lead to the emission of light, a phenomenon called sonoluminescence (SL). The study of sonoluminescence has led to refinements in understanding cavitation. This talk will focus on three areas of research in which the application of sonoluminescence is used to investigate cavitation: (1) In the field of single-bubble SL, techniques have been improved to study the dynamical motion of the bubble and, with comparisons to SL, models of the bubble dynamics have been refined. In particular, thermal conduction and the trapping of water vapor during the bubble collapse contribute significantly to the internal temperatures and to the dynamics of the cavitation bubble. (2) In lithotripsy, intense shock waves generate cavitation fields which play a role in kidney stone communition. *In vitro* studies have shown that SL can be correlated with a bubble's expansion and collapse within the field. (3) In sonochemistry, high-intensity ultrasound is used to initiate or enhance chemical reactions. SL is often observed under these conditions, and in some cases, there is a correspondence between SL intensity and sonochemical yield. Although often overlooked, it can be important to account for the physics and chemistry at the bubble surface.

9:40–9:55 Break

9:55

2aBB6. Microcavitation-assisted selective surface erosion. Sameer I. Madanshetty (Mech. and Nuclear Eng., Kansas State Univ., 322 Rathbone Hall, Manhattan, KS 66506)

Cavitation is a chance-dominated process and cavitation activity cannot be readily controlled. Interestingly, our recent research in ACIM (acoustic coaxing induced microcavitation) allows us to bring about microcavitation quite readily at all surfaces of interest. Liquid-borne microparticles exposed to near threshold strength sound fields do not ordinarily cause cavitation, however, experience with ACIM indicates that cavitation by the microparticles is readily facilitated and one can bring about controlled erosion of painted and inked surfaces. Cavitation is an energy concentrator and microcavitation implosions concentrate enormous energy at points. ACIM, consequently, can generate high-energy density points on liquid-borne surfaces. Since the implosion effects are felt only locally, these high-energy density points easily remove ink/paint off surfaces while preserving the structural integrity of the underlying

substrate. The presentation will demonstrate initial evidence of using silent sound and clean water-controlled acoustic microcavitation for removing xerographic ink from paper, and paint layers from metal surfaces without damaging the underlying substrates, paper and metal, respectively. More significantly such microcavitation-assisted surface erosion permits an easy method for determining the bonding strength between the thin film and its substrate. [Work supported by Navy and NSF.]

Contributed Papers

10:20

2aBB7. Surf-zone bubble detection using multiple techniques: The Worbarrow Bay experiment. Timothy G. Leighton, Matthew D. Simpson, Steve D. Meers, Paul R. White (Inst. of Sound and Vib. Res., Univ. of Southampton, Highfield, Southampton SO17 1BJ, UK), Gary J. Heald, Hugh A. Dumbrell, James W. Clarke (DERA Bincleaves, Weymouth, Dorset, UK), Peter R. Birkin, and Yvonne Watson (Univ. of Southampton, Highfield, Southampton SO17 1BJ, UK)

This paper describes a multisensor experiment to characterize the bubble population and its effects in the surf zone. Active bubble detectors (combination-frequency sensors, and the inversion of sound speed and attenuation) provided estimations of the bubble-size distribution, and allowed better interpretation of more novel sensors. These include acousto-electrochemical sensors, and estimations of the bubble population from the ring-up time associated with the insonification of the population, and from the ambient noise (attention being paid to how the ringing population so determined differs from the full population measured by active techniques). [Work supported by the U.K. Defence Evaluation Research Agency and the EPSRC.]

10:35

2aBB8. The effects of oceanic surfactants on acoustic propagation in bubbly liquids. Joseph C. Jankovsky and Ronald A. Roy (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, jankov@bu.edu)

The chemical composition of the top layer of the ocean is known to contain surface-active substances that are readily adsorbed to an air-water interface. The presence of surfactants in a liquid induces a viscoelastic stress along a two-dimensional interface. These surfactants can coat bubbles that alter their individual dynamics. A model is presented that incorporates the effects of surface viscoelasticity for acoustic propagation in bubbly fluids. The effects of both surface dilatational viscosity and surface elasticity on the phase speed and attenuation are considered. Surface viscosity is found to decrease the attenuation near bubble resonance frequencies, yet increase damping below resonance. Surface viscosity also diminishes the resonant effects for the phase speed in bubbly fluids. Both effects become significant for bubble sizes below 100 microns. The addition of surface elasticity is found to decrease the mean oscillation bubble radius, and thus shift bubble resonance to a higher frequency. The effects of the model on *in situ* acoustic bubble sizing methods will also be discussed. [Work supported by ONR.]

10:50

2aBB9. Acoustic scattering from an elastic tube filled with bubbly fluid. Preston S. Wilson, Ronald A. Roy, and William M. Carey (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215)

A complete model describing broadband sea surface scattering at high wind speeds has not been developed. One difficulty is accounting for scattering from near-surface bubble clouds. This problem has been addressed in the literature for low frequencies. To first order, an acoustically compact bubble cloud can be modeled as a compressible sphere, where the scattering strength depends only on spherical cloud volume and mean void fraction, not the bubble size distribution or cloud shape. This hypothesis has been experimentally tested using freely rising artificial bubble clouds [J. Acoust. Soc. Am. **92**, 2993–2996 (1992)]. The measured low-frequency monopole target strength of the cloud agreed with theory but higher-frequency results did not. To further understand scattering from these objects, laboratory scattering experiments are underway using geometrically well-characterized bubbly fluid targets. Initial measurements of scattering from a bubbly fluid-filled latex tube are presented and compared to an effective medium theory. These initial results lack independent void fraction determination but good qualitative agreement is found, even above the monopole resonance frequency. A new method used in these experiments to generate large volumes of nearly monodisperse bubbly fluid samples will also be described. [Work supported by ONR.]

11:05

2aBB10. Acoustic scattering from partially voided compliant and fluid spheres. Joseph C. Jankovsky, Ryan D. McCormick, Ronald A. Roy, and William M. Carey (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, jankov@bu.edu)

The presence of bubbles has been shown to change the compressibility and complex sound speed in a liquid. In the ocean, acoustically compact bubbly mixtures manifest themselves as highly compressible regions that effectively scatter low-frequency sound. To study low-frequency sound scattering, multifrequency backscattering experiments have been performed in a tank using partially voided three-quarter-inch diameter polyurethane spheres as targets. Target strengths (2-20 kHz) were measured for four spheres with void fractions of 0%, 3.4%, 4.2%, and 6%. Measured target strengths for the voided spheres were on the order of -40 to -60 dB (re 1 m). The frequency response exhibited modal structure, with peaks shifting to lower frequencies for higher void fractions. No backscatter signal was detected for the solid polyurethane sphere. Target strength was also measured for a hollow polyurethane sphere containing a suspension of bubbles in polymer gel. The void fraction was determined by fitting the scattering theory and low-frequency bubbly fluid compressibility model to the measured data. [Work supported by ONR.]

11:20

2aBB11. Comparing the predictions of a numerical model of SBSL in a variable acceleration environment to experiment. Charles Thomas, Sean Wyatt,[a] Ronald Roy, and R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215)

The results of an August 1999 KC-135 SL experiment will be discussed and compared to a numerical model [Wyatt *et al.*, J. Acoust. Soc. Am. **106**, No. 4, Pt. 2, 2290 (1999)] for SL in a cubic acoustic resonator in a varying acceleration environment. The model takes into account gravitational effects and varying ambient pressure effects, and predicts the maximum bubble size, light intensity, and levitation position of the bubble. A review of and comparison to other investigators experimental results will be included. [Work supported by NASA.] [a] Currently at Ford Motor Company, Detroit, MI.

11:35

2aBB12. Sonoluminescence at the nanoscale. Carlos G. Camara, Keith R. Weninger, and Seth J. Putterman (Phys. Dept., Univ. of California, Los Angeles, Los Angeles, CA 90095)

Sonoluminescence (SL) has been observed to be robust over a wide parameter space ranging from 8 kHz to 11 MHz. Although some lines can be discerned in the spectra of some cavitation clouds, they sit on top of a broadband ultraviolet continuum that spans to at least 6 eV. Experiments and theory indicate that the light-emitting region can reach tenths of nanometers. Although the SL mechanism and its huge parameter space remain a mystery, it has already been put to use as a surgical device. At 30 kHz it is used for internal lipectomy and at 1 MHz it is used for externally assisted lipectomy. At 11 MHz a dense cloud of light-emitting bubbles is observed in the far field of the transducer. These observations could be exploited for future uses of nonintrusive ultrasound surgery.

11:50

2aBB13. Anomalous changes of sonochemical luminescence intensity under stirring operation. Hideto Mitome (Natl. Industrial Res. Inst. of Nagoya, AIST, Nagoya, 462-8510 Japan, mitome@nirin.go.jp) and Shinichi Hatanaka (Japan Sci. and Tech. Corp., Kawaguchi, Saitama, 332-0012 Japan)

Measurement of sonoluminescence is an effective method to evaluate a reaction field induced by intense ultrasound in a liquid. The present paper discusses changes of intensity of sonochemical luminescence from a luminol solution under various stirring conditions. A glass beaker filled with a luminol solution was irradiated by ultrasound using Langevin-type transducers attached to the bottom and side walls of a cooling bath. As in conventional chemical reactions, the intensity of luminescence was expected to increase with stirring operation. At 28 and 100 kHz, the luminescence intensity increased at a higher applied power and at a faster stirring speed. But, at 45 kHz where intense cavitation was observed in the present apparatus, the luminescence intensity decreased as the applied power increased while stirring with a flat paddle. This anomalous behavior was more pronounced at more intense stirring operation with an impeller. The authors reported the existence of upper threshold of sound pressure for sonoluminescence [Hatanaka *et al.*, Jpn. J. Appl. Phys. **38**, 3053–3057 (1999)]. The present results suggest that stirring operation affects the generation and growth of cavitation bubbles making them exceed the upper threshold in some cases, and that there is an optimum stirring motion for sonochemical reactions.

TUESDAY MORNING, 5 DECEMBER 2000 PACIFIC SALON A, 8:35 TO 11:15 A.M.

Session 2aEA

Engineering Acoustics: Ultrasonic Sensors and Motors

Harold C. Robinson, Chair
Naval Surface Warfare Center, Code 2131, 1176 Howell Street, Newport, Rhode Island 02841-1708

Chair's Introduction—8:35

Invited Papers

8:40

2aEA1. Ultrasonic motors (USM)—An emerging actuation technology for planetary applications. Yoseph Bar-Cohen, Xiaoqi Bao, and Hari Das (JPL/Caltech, MC 82-105, 4800 Oak Grove Dr., Pasadena, CA 91109-8099, yosi@jpl.nasa.gov)

Efficient miniature actuators that are light, compact, and driven by low power are needed to drive telerobotic devices and space mechanisms in future NASA missions. Examples of space mechanisms and devices that require actuators include robotic arms, miniature rovers, release mechanisms, positioning devices, aperture opening and closing devices, and real-time compensation for thermal expansion in space structures. These motors need to operate at various temperatures and pressures with a large range of thermal variations over a relatively short period swing. Ultrasonic rotary motors have the potential to meet this NASA need and they were developed as actuators for miniature telerobotic applications. These motors were adapted for operation in the environment of Mars, which includes very low temperatures and vacuum. A hybrid analytical model, including the influence of the rotor and stator dynamics, friction effects, and interface effects, was developed to design an efficient ultrasonic motor as a complete system. In parallel, efforts have been made to determine the thermal and vacuum performance of these motors, and effective operation at temperatures as low as $-150\,^{\circ}\mathrm{C}$ and at a pressure of 16 mtorr were demonstrated. To explore telerobotic applications for USMs a robotic arm was constructed with such motors.

9:05

2aEA2. Bio-inspired acoustic sensors based on artificial stereocilia. Flavio Noca, Michael Hoenk, Brian Hunt (Jet Propulsion Lab., 4800 Oak Grove Dr., Pasadena, CA 91109, flavio.noca@jpl.nasa.gov), Petros Koumoutsakos, Jens Walther, and Thomas Werder (ETH Zurich, Zurich, Switzerland)

A unique, biologically inspired acoustic transducer based on artificial stereocilia is being developed. This transducer will enable directional sensitivity and miniaturization of acoustic sensors while enhancing sensitivity, ultimately leading to revolutionary advances in acoustic detection and signal processing. The similarities between natural stereocilia and the proposed transducer array could allow for the first time the fabrication of an artificial cochlea that relies on biologically inspired signal-processing techniques. Other applications, such as measuring sounds generated by moving micro-organisms and nanoscale biological events (metabolic flows), may be enabled by this novel transducer. Finally, artificial stereocilia (functioning as actuators instead of sensors, analogous to stridulatory pegs in insects) will be capable of generating acoustic signals for applications in active acoustic instruments such as sonar. Before these advantages can be realized, fundamental developments are required in the design and fabrication of a transducer based on artificial stereocilia arrays. The potentially revolutionary nature of this transducer is being demonstrated by fabricating and characterizing acoustic sensors based on recently produced carbon nanotube arrays.

9:30

2aEA3. A surface acoustic wave hygrometer for high-resolution measurements of atmospheric humidity. Michael E. Hoenk, Greg Cardell, Flavio Noca, and Robert K. Watson (Jet Propulsion Lab., California Inst. of Technol., Pasadena, CA 91109)

Weather and climate depend on atmospheric phenomena at all scales, from large-scale atmospheric circulation to microphysical processes. In this talk, a surface acoustic wave (SAW) hygrometer developed at JPL for high-resolution measurements of atmospheric humidity will be discussed. Flight tests of an early prototype on the NASA DC8 showed more than an order of magnitude faster response than chilled-mirror hygrometers. In 1998, a 1-kg reference radiosonde based on the SAW hygrometer was flown on a small balloon to an altitude of 44 000 feet, while a mobile ground station recorded telemetered data containing high-resolution measurements of humidity, pressure, temperature, and GPS position. Later in 1998, the SAW hygrometer was deployed on the NASA DC8 to measure humidity in Atlantic hurricanes. In 1999, the SAW hygrometer recorded humidity in the test flight of NASA's Helios unpiloted aircraft. Currently, this instrument is being modified for incorporation into a hand-held instrument for monitoring environmental humidity on the space station. [The research described in this paper was performed at the Jet Propulsion Laboratory, California Institute of Technology, and was jointly sponsored by the National Aeronautics and Space Administration, Earth Science Enterprise, Space Science Enterprise, and Human Exploration and Development of Space Enterprise.]

9:55–10:10 Break

10:10

2aEA4. The zNose, a new electronic nose using acoustic technology. Edward J. Staples (Electronic Sensor Technology, LP, 1077 Business Ctr. Circle, Newbury Park, CA, staples@estcal.com)

An array of sensors simulating the human olfactory response has become known as an electronic nose, or eNose. An eNose provides a vectorial image in N-dimensional space (where N equals the number of sensors) of specific vapor mixtures containing possibly hundreds of different chemical species. Typically, eNoses have only a few sensors, produce responses that are not correlated, and have poor sensitivity. A new type of electronic nose, based upon fast chromatography and a single high-Q acoustic sensor, solves these problems by stimulating a virtual sensor array containing hundreds of orthogonal sensors. Analysis of any odor is accomplished by serially polling this virtual sensor array or a spectrum of retention times. This paper will describe the acoustic technology as well as present results for applications involving important vapors associated with several commercial foods and beverages. Part per billion sensitivity has been achieved with volatile compounds and part per trillion sensitivity for semi-volatile compounds. Because the new acoustic technology is quantitative it is the only electronic nose technology to be validated by the US EPA. The commercial expression of this technology, the zNose, is now providing an on-line quantitative measure of quality for food, beverages, cosmetics, and other aromatic products.

10:35

2aEA5. Recent developments in transducers: Opportunities for the future. Ilene Busch-Vishniac (Whiting School of Eng., Johns Hopkins Univ., NEB 120, 3400 N. Charles St., Baltimore, MD 21218)

Most innovations in sensors and actuators result either from the introduction of new materials, or the exploitation of newly discovered mechanisms for conversion of energy between the electrical and mechanical domains. This talk reviews new recent transducer improvements and suggests opportunities for the future of ultrasonic sensors and actuators. Although many of the transducer examples portrayed were not conceived for ultrasonic or acoustic applications, the exploitation of their materials and energy conversion mechanisms could provide a means of enhancing the arsenal of ultrasonic transducers available.

Contributed Paper

11:00

2aEA6. MEMS condenser microphone. Peter J. Henning and Torben Storgaard-Larsen (Bruel & Kjaer, 2850 Naerum, Denmark)

The measurement microphones of today are the results of almost 50 years of continuous optimization of production methods based on precision mechanics. Scientists have searched for new ways of measuring sound, which can reach the performance of the condenser microphone, but so far have not succeeded! However, applying new materials and production methods, the window of achievable specifications of condenser microphones can be widened. Combining the accuracy and reliability of silicon micro machining with the experience gained on traditional measurement microphones, unique new features can be added to the range of microphones. As a first result of such efforts, a new microphone unit, the multifield microphone, has been developed. The multifield microphone has a ''flat'' frequency response in all sound fields, i.e., pressure field, diffuse field, or free field (when the sound comes from different angles of incidence). This results in higher accuracy in most measurement situations, where the sound field is nonidealized.

Session 2aNSa

Noise and NOISE-CON: Soundscapes

Brigitte Schulte-Fortkamp, Cochair
Department of Physics/Acoustics, Oldenburg University, Oldenburg D-26111, Germany

Joseph Pope, Cochair
Pope Engineering Company, P.O. Box 590236, Newton Centre, Massachusetts 02459-0002

Chair's Introduction—8:40

Invited Papers

8:45

2aNSa1. Can soundscapes work as a moderator concerning noise annoyance? Brigitte Schulte-Fortkamp (Dept. of Phys./Acoust., Oldenburg Univ., 26111 Oldenburg, Germany, brigitte@aku.physik.uni-oldenburg.de)

The evaluation procedure to detect and to analyze soundscapes in their meaning regarding noise annoyance measurement is a new step and a very complex procedure. Focusing on the interaction of people and noise, different aspects like the structure of urban areas, people living in those areas, architectural and social parameters designing those areas, and acoustical and visual parameters will be taken into account for analyses. Moreover socio-cultural- and lifestyle-related elements of soundscapes have to be discussed concerning the aspects of annoyance measurement which usually are focused on interference, disturbance, sources of annoyance, and predicting noise annoyance. Basically, in the discussion of both procedures it is questioned whether soundscapes can work as a moderator concerning annoyance. Detailed aspects may lead to the question of correlations between annoyance judgments and somatic processes, and what happens when people give an overall judgment on the degree of annoyance. Future research is needed to explore soundscapes moderating the judgment on life situation and noise sources.

9:05

2aNSa2. Soundscapes in American mainstream films. Barbara Flueckiger (Inst. of Film Theory, Univ. of Zurich, Plattenstrasse 54, CH-8032 Zurich, Switzerland, zauberklang@bluewin.ch)

The question of how films establish an acoustic environment was studied in the course of a research project on sound design in American mainstream film. The study was supported by the Swiss National Foundation. The corpus of this research project consisted of 96 films, their production years ranging from 1926 to 1995. Most of these films received an Academy Award for Best Sound. They were closely examined regarding their strategies to establish fictional, yet natural-seeming, soundscapes. It was discovered that the film industry developed a rather restricted vocabulary to display certain types of geographically, socially, and/or culturally defined places. The cause of the restriction can be found in technical and historical reasons as well as psychological considerations. Film soundscapes have a clear communicative function in contrary to natural soundscapes, which contain random noises. According to an unwritten Hollywood rule of the so-called classic era of the thirties and forties, any distraction from the narrative goal had to be avoided. This rule, however, was revised when multichannel formats were developed, first in the fifties with different wide-screen systems, and later in the mid-seventies with the Dolby stereo 4- to 6-channel sound system.

9:25

2aNSa3. Sound quality of exterior vehicle noise. Klaus Genuit (HEAD Acoustics GmbH, Eberstrasse 30a, 52134 Herzogenrath, Germany)

For several years the acoustical engineers from all car manufacturers have been busy investigating and improving the sound quality of the interior noise of vehicles. For 20 years we have been using specific measurement technologies and analyses to describe sound quality with consideration of the signal processing of human hearing. But with respect to the exterior vehicle noise only standard measurements, e.g., the A-weighted sound pressure level, are still in use. Although the A-weighted sound pressure level has been reduced, the annoying effect of exterior vehicle noise and traffic noise has increased. The negative impact of exterior vehicle noise and traffic noise on man must be reduced. Solutions must therefore be found to the vehicle noise emission problem while allowing sustainable mobility. An important prerequisite for improvement of the well-being of the public and acceptance of future traffic and acoustic environments is the consideration of human sound perception of the exterior vehicle. This paper describes how we can transfer the knowledge of sound quality analysis with respect to interior noise to the evaluation of exterior vehicle noise. This includes not only the methodologies for the measurement and analyses, but also the development of a prediction tool.

9:45

2aNSa4. Soundscapegraphy of old-town of Kyoto with the Gion Festival. Kozo Hiramatsu (Mukogawa Women's Univ., 663-8558 Nishinomiya, Japan, NBG01036@nifty.ne.jp)

The soundscape of the Yamahoko-cho area in the old capital of Japan, Kyoto, is described from various aspects based on an intensive survey of acoustic ecology. The area located in the city center is famous for the Gion Festival, which is one of the biggest and oldest festivals in the country. The area's soundscape on ordinary days is basically dominated by the noises of automobiles passing through the streets, since there are also companies and shops in the area, which is the trading center of Japanese costumes, kimono. In July, when the ceremony and festival are undertaken for 1 month, the soundscape changes to be dominated by a variety of sounds related to the festival. Particularly during the days when bands of music by flutes, drums, and small gongs are performed, the soundscape of the area becomes a complete festival one. The music is so famous that it is recognized as characteristic of the city. The residents in the area are raised listening to the sounds in summer, and these sounds are closely connected with their lives. Though the music is no doubt the loudest sound heard in the area, no one complains.

10:05

2aNSa5. Factors moderating the effect of noise barriers. Tommaso Meloni and Fredy Fischer (Swiss Agency for the Environment, Forests and Landscape, 3003 Berne, Switzerland, tommaso.meloni@buwal.admin.ch)

More often than not, on noise research the dose-response relationship shows a modest explained variance. A field study was carried out to investigate the effects of 13 different noise barriers, and 500 people exposed to noise were surveyed using questionnaires. First it seems that the acoustic effectiveness of the noise barriers is responsible for the reduction in annoyance. However, a detailed analysis shows that further factors moderate the relation between the noise parameters and the reaction to noise. It seems that simply introducing a noise barrier can cause a reduction in annoyance to an extent equivalent to 5 dB. This confirms the fact that the outdoor noise levels and the acoustic design of the surroundings are among the principle parameters influencing annoyance. In addition, annoyance is determined by individual acceptance of noise barriers. The degree of acceptance splits the reaction curve with a separation around 4 units on the rating-scale of 11 units of annoyance. The design of the noise barrier partly affects this splitting. There is a higher degree of acceptance of naturally constructed and transparent noise barriers than for ones that are specially formed. Other determining factors for annoyance will be discussed.

10:25–10:45 Break

Contributed Papers

10:45

2aNSa6. Quiet areas and quiet facades: Important elements in the EU Noise policy. Tor Kihlman (Dept. of Appl. Acoust., Chalmers Univ. of Technol., SE 412 96 Goteborg, Sweden)

The long-term goal for road traffic noise, Lden$<$55 dB, cannot be reached for all dwellings either in sprawled or in compact cities during the foreseeable future [T. Kihlman and W. Kropp, ''Limits to the Noise Limits,'' ICA, Seattle 1998]. What, then, can be done to improve the situation? One strategy is to exploit the spatial noise level variations. Even though many dwellings are exposed to Lden=65–70 dB, outdoor levels in the range 40–50 Lden are not rare in the cities: courtyards in European cities being the typical case. The EU draft directive on ambient noise puts emphasis on the protection of quietness. Quiet areas and ''quiet facades'' shall be shown on noise maps. The concept of a relatively quiet facade has been introduced. Questions to answer are: What are the effects on the annoyance and sleep disturbance when dwellings have one noisy and one quiet side? To what extent do existing, noise-exposed dwellings also have a quiet side? What are the potentials to increase also at low cost—the access to quietness in the urban renewal process? Some results from ongoing research will be presented as well as plans for further research.

11:00

2aNSa7. Measurement of the natural soundscape in national parks. J. Micah Downing (Wyle Labs., 2001 Jefferson Davis Hwy., Ste. 701, Arlington, VA 22202, mdowning@arl.wylelabs.com) and Eric Stusnick (San Jose, CA 95135)

The National Park Service (NPS) has identified the natural soundscape in its various park units as one of the resources it has been mandated to protect. Natural soundscape is defined as the acoustic environment that would exist in the absence of human-related activity. Accordingly, the NPS has developed policies related to soundscape management, preservation, and restoration, which require information about the natural ambient sound levels in all of their properties throughout the country. As part of this process, metrics and procedures for measuring the natural soundscape and assessing non-natural intrusions have been developed. Two general types of procedures have been developed: one based on audibility considerations, the other based on acoustic energy measurements. This paper describes the work that has been carried out in developing and assessing these procedures.

11:15

2aNSa8. Measurement and modeling of snowmobile noise and audibility at Yellowstone and Grand Teton National Parks. Christopher Menge and Jason Ross (Harris Miller Miller and Hanson, Inc., 15 New England Executive Park, Burlington, MA 01803)

The U.S. National Park Service is concerned about the effect that noise from snowmobiles and other over-snow vehicles has on the natural soundscapes in the National Parks. This paper addresses the measurements and modeling performed to assess both the total area in the parks where snowmobile noise would be audible and the noise intrusiveness as a function of distance from the trails. The work was performed in support of an Environmental Impact Statement that evaluated seven operational transportation alternatives in the two parks. The modeling approach computed both A-weighted sound levels and audibility continuous time histories of snow-vehicle noise at several distances from the trails. Acoustical input to the model included $\frac{1}{3}$-oct band levels for ambient conditions, vehicle source levels, and propagation of sound over a snow-covered surface. Consequently, distances to the onset of audibility detection, areas of the parks affected, and statistics on received sound levels were calculated. The paper presents both the modeling approach and examples of results.

11:30

2aNSa9. The application of a noise-rating scale composed of hue-scale and L_{Aeq} on various kinds of noise sources for noise map. Kenji Furihata and Takesaburo Yanagisawa (Dept. of Elec. and Electron. Eng., Faculty of Eng., Shinshu Univ., 4-17-1 Wakasato, Nagano, 380-8553 Japan, kennfur@gipwc.shinshu-u.ac.jp)

In addition to reporting the results of measurements of existing environmental noise and the results of calculations of noise from projected activities, the application of color for representing in terms of noise zones may be useful (ISO 1996-2). Between color and noise, in spite of differences, there exists an interaction, although the connection between the visible and the audible expression is subjective. Subjectivity can in a given moment become objective, if one succeeds in unifying the visible and the audible. In the case of colored paper on a white background, the relation between the hue level expressing the psychological effect (annoyance) and $L_{Aeq(5min)}$ (dB) of typical noise sources was discussed in a laboratory experiment. According to the obtained noise-rating scale, the seven-hue scale and $L_{Aeq(5min)}$ are correlated well ($r=0.874$) as follows: white (value V9; not at all bothersome) ⇔ below 26 dB, light blue (hue 2.5B, value 8/chrome 4; not bothersome) ⇔ 26 to 37 dB, light green (2.5BG, 7/8; not too bothersome) ⇔ 37 to 49 dB, bright yellow (7.5YR, 7/14; a little annoying) ⇔ 49 to 61 dB, brown (10R, 4/8; annoying) ⇔ 61 to 70 dB, purple (2.5P, 4/10; very annoying) ⇔ 70 to 80 dB and red (5R, 5/14; extremely annoying) ⇔ over 80 dB.

11:45

2aNSa10. Methodology for analyzing perceptual criteria of urban soundscape. Raimbault Manon (CERMA UMR CNRS 1563 Rue Massenet, 44000 Nantes, France)

The exposed study is part of a larger program which intends to constitute an urban soundscape simulator. This paper presents analyses of urban scenes with the aim of identifying parameters of sound perception for this future model. It seems that the perception of everyday urban sounds is not yet well understood by urban planning actors. However, those considerations are more and more problematic in everyday life and are becoming increasingly important for town planner recommendations. Therefore the assessment of perception indicators linked to acoustical measures is intended to help one fully understand urban acoustics phenomena. The method consisted of two simultaneous procedures on site. In order to calculate acoustical parameters, we recorded samples of soundscape that we analyzed afterwards. At the same time we observed perceptual attitudes and opinions using a semantic differential inquiry of passers-by. The locations were chosen in order to represent the main places of European city centers. Experimental results showed that five sector-related indicators could be defined. A full understanding of urban acoustical phenomena needs this multidisciplinary approach so that both physical and psychosociological parameters can be integrated in town planning.

TUESDAY MORNING, 5 DECEMBER 2000 SCHOONER ROOM, 7:55 TO 11:55 A.M.

Session 2aNSb

Noise and NOISE-CON: Power Plant Noise Control and Prediction and Industrial Noise

Frank H. Brittain, Chair

Bechtel Corporation, 50 Beale Street, San Francisco, California 94105

Chair's Introduction—7:55

Invited Papers

8:00

2aNSb1. Noise control engineering challenges in designing power plants to meet noise limits. Robert Putnam (Siemens Westinghouse, 4400 Alafaya Trail, Orlando, FL 32826)

Noise control considerations play a vital and often pivotal role in a broad range of issues in the design of modern power plant facilities. The entire design process, from site selection, through final operational acceptance, is affected by the proper integration of the governing noise criteria, especially when relatively low levels are called for. The noise control engineer is continually challenged to justify added noise control measures, maintain adequate design margins, minimize costs, performance impacts, and design impacts, while coordinating a number of cross-discipline interfaces ranging from marketing and sales to contracts to startup and field engineering. This paper presents an overview of the principal noise control engineering tasks involved when effectively integrating required noise control features into the design process, highlights key interface requirements, and provides notes on the practicality, safety, and maintainability of control features. Illustrative examples from the author's experience, both positive and negative, will be included.

8:20

2aNSb2. Using a prediction model to allocate allowable noise between sources and establish equipment noise limits. David J. Parzych (Power Acoustics, Inc., 12472 Lake Underhill Rd. #302, Orlando, FL 32828, dparzych@poweracoustics.com)

Power generation facilities typically require customized noise abatement features to achieve compliance with various local and state noise regulations. Differing equipment or equipment arrangements, size and placement of equipment on the plant owner's property, and location within a community can all affect the amount of noise control necessary for a given facility. Selecting the correct amount of silencing for each piece of plant equipment is essential when optimizing for reduced cost of the overall plant noise control treatments. To achieve the overall plant noise goals, an analytical noise prediction model of the facility is built and exercised with various equipment options until the desired goal is achieved. The paper describes methods of using a noise prediction model to

allocate noise ''budgets'' among sources, and using the model to establish noise criteria in specification form. The discussion will address the overall effects of starting with poorly defined data, as well as imprecise or erroneous noise source data supplied by equipment manufacturers. Estimating reasonable and achievable noise goals for the equipment will be addressed.

8:40

2aNSb3. Integrating noise controls into the power plant design. Wayne E. Bradley (Stone & Webster Eng. Corp., 245 Summer St., Boston, MA 02210, wayne.bradley@stoneweb.com)

The selection of noise controls and the specification of low-noise equipment is an important part of designing a power plant to meet specific noise limits. However, successfully interfacing with the design team and integrating these controls into the project design is an equally demanding task. The noise-control specialist must work closely with the design team to assure that the affects of the mitigation on space allocation, equipment performance, energy consumption, cooling, access to equipment, and safety are appropriately considered. Each of these factors must be addressed to assure that the controls work effectively under plant operating conditions. The design process is iterative—design details and controls may change as the design progresses. Both formal and informal lines of communication must be cultivated with the design team, and the mitigation recommendations made to them should be followed up to assure their proper incorporation into the project design. This paper describes these processes.

9:00

2aNSb4. Noise-control engineering interacting with the design process of a power plant. Marlund E. Hale (Parsons Eng. Sci., Inc., 100 W. Walnut St., Pasadena, CA 91124)

There are many challenges which face the noise-control engineer associated with developing and implementing both the occupational and community noise controls for power generation facilities. They include becoming an early and viable part of the proposal and design process and maintaining a proper and profitable business effort. Challenges during the proposal and design phase include working within the business plan, dealing with cost estimates incorporated during the development of the winning proposal budget, the project design schedule, the engineering process, and the noise-control cost estimating and procurement process. Once a project has been won, certain financial and scheduling commitments are already in place. The project design team is then assembled and the design evolves conceptually as the various issues affecting the project are incorporated. To be a viable and effective part of the design team, the noise-control engineer must become an early player, deal with an incomplete and evolving design, keep communication lines open, stay abreast of the many facets of the design evolution, communicate project-impacting information to project management, interface with the procurement process, and stay in a position to interface with and support engineering and construction inspectors.

9:20

2aNSb5. Risks in designing power plants to meet noise limits and how to manage them. Frank Brittain (Bechtel Corp., 50 Beale St., San Francisco, CA 94105)

Risks and risk avoidance affect both strategy and many choices made in designing power plants to meet noise limits. Each player typically attempts to avoid risk by passing off as many as feasible to other players. Many of these risks can readily be anticipated, but others are often unexpected. Each player—noise control engineer, engineer-constructor, owner, banker, and regulator—has somewhat different risks. Many factors can affect risks, including technical uncertainties, noise criteria, unreliable noise source data, regulatory environment, limits to the state-of-the-art, costs, schedule, any lack of experience of the responsible noise control engineer, and equipment that does not meet its noise limits. Risks are greatly increased when liquidated damages are present, the responsible noise control engineer is not adequately supported, and noise control is neglected or initiated too late in the design process. All of these factors are interrelated in complex ways. To successfully design a power plant, risks and risk avoidance should be understood, and strategies developed to reduce those risks. This paper builds on previous papers in the Special Session, reviews the risks of each player, and suggests strategies to reduce risks. Further, the ways in which risk affect the design process are identified.

9:40

2aNSb6. Strategy for writing procurement specifications. George F. Hessler, Jr. (Hessler Assoc., Inc., 3862 Clifton Manor Pl., Haymarket, VA 20169, gfhhai@msn.com)

The key word for writing workable procurement specifications is clarity. The recipient is much more interested in responding when the specification requirements are understood and unambiguous. As a minimum, the procurement specification must describe the starting point and end result specified by the author. The method of test verification and allowable tolerances, if any, is critical and must be defined in the procurement specification. Current acoustical requirements for power plant equipment must be specified in the near and far field with a clear order of precedence if there is a conflict between the two locations. The trend today for specifying far-field sound levels is to limit the specified parameter to the overall ''A''-weighted sound level. Since this parameter offers no chance of preventing excessive low-frequency noise, the ''C''-weighted limit, as a minimum, must also be concurrently specified for all types of power plants.

10:00

2aNSb7. Customer sound criteria and vendor acoustical performance guarantees: The difficulties of meeting the first and verifying the second. Robert S. Johnson, Sr. (Solar Turbines, Inc., 9280 Sky Park Court, San Diego, CA 92131)

In their procurement specifications, gas turbine end users include acoustical criteria. Gas turbine manufacturers, in accepting the criteria, specify the performance of acoustical components which are procured from acoustical component vendors for installation on the sound sources supplied by the gas turbine manufacturer. If site noise problems develop, the sound sources must be investigated

and the performances of the acoustical components evaluated to determine if the performance guarantees have been met. This paper discusses one gas turbine manufacturer's experience with how sound criteria are specified and responded to. Also discussed, are problems associated with the verification of vendor product performance guarantees, and the resolution of perceived product performance deficiencies.

10:20–10:40 Break

Contributed Papers

10:40

2aNSb8. Active control of low-frequency turbine exhaust noise. Bruce Walker, Alan Hersh, Joseph Celano (Hersh Walker Acoust., 780 Lakefield Rd., Unit G, Westlake Village, CA 91361), and Eldon Ray (Braden Manufacturing, Tulsa, OK 74101)

Gas turbine generators radiate noise from their exhaust stacks that contains substantial acoustical energy at very low frequencies. This kind of noise can result in community annoyance due to window rattling and other structural vibration excitation as well as disturb residents. Because of the long wavelengths of sound at these frequencies, typical duct sound absorbing treatments are inefficient, large and expensive. A research program was conducted to demonstrate the ability of a Helmholtz-resonator-based active noise control system to attenuate low-frequency noise in a heated exhaust stack. A unique thermal stratification concept was developed that successfully protected the sound sources from hot exhaust flue gases. The thermal stratification concept permitted the use of both conventional off-the-shelf JBL 18-in. electrodynamic actuators and custom-fabricated piezoceramic sound sources in the demonstration program. Test results showed that 14–22 dB of noise attenuation was achieved over the frequency range 20–32 Hz at flue temperatures between 80 ° and 1300 °F.

10:55

2aNSb9. Certifying noise emissions from heat recovery steam generators (HRSG) in complex power plant environments. George F. Hessler, Jr. (Hessler Assoc., Inc., 3862 Clifton Manor Pl., Haymarket, VA 20169, gfhhai@msn.com)

HRSG units are used to generate steam by capture of heat from the combustion turbine exhaust gas, typically at 1000 °F. Combustion turbines, nominally in the 100–200-MW size, generate approximately 150 dB, *re* 1 pW ''A'' weighted sound power. Hence, noise emissions from the HRSG wall surfaces and stack outlet are a significant source of power plant noise in today's combustion turbine-based plants. Noise mitigation measures are required for most installations, and it is the responsibility of the HRSG supplier to engineer, furnish, and warrant the performance of the HRSG assembly. Verification of HRSG noise emissions in the field is difficult since the units are typically surrounded by countless other plant noise sources in complex arrangements. There is no applicable test standard in place for reference. This paper presents alternative methods of isolating and verifying HRSG noise source emissions in complex environments, and recommends test methods based on experience.

11:10

2aNSb10. Determination of sound pressure levels *in situ* using sound intensity measurements. Stephen E. Keith, Vincent Chiu (Health Canada, Radiation Protection Bureau, 775 Brookfield Rd., 6301B, Ottawa, ON K1A 1C1, Canada, skeith@hpb.hwc.ca), and G. Krishnappa (Natl. Res. Council of Canada, Ottawa, ON K1A 0R6, Canada)

Measurements of emission sound pressure levels of machinery require either specially defined test rooms or calculated corrections for the acoustic environment. In principle, it is possible to determine emission sound pressure levels from sound intensity measurements at specified work stations in any test environment if the requirements of background noise levels and field indicators are fulfilled. The draft international standard ISO 11205/CD specifies such a method. In this paper the accuracy of emission sound pressure levels using sound intensity measurements was examined for three small sources in three acoustic environments, an anechoic environment with loudspeakers to simulate background noise, an office environment, and a reverberant environment inside a stairwell. In the first two environments good measurement accuracies, within 1 dB, were obtained. Sound intensity measurements by pointing the probe towards the source were as accurate, and simpler than computation of the resultant intensity using three arbitrary orthogonal measurements. As predicted by field indicators, measurements in the stairwell gave unacceptable errors for all three sources.

11:25

2aNSb11. Diagnostics of vibration excitation mechanisms for complex hydromechanical systems. Andrei B. Prokofiev, Valeri V. Lenchine, and Evgeniy V. Shakhmatov (Inst. of Acoust. of Machines, Samara St. Aerosp. Univ., Moskovskoe shosse 34, Samara 443086, Russia)

It is known that high-amplitude vibration can significantly impair reliability of technical tools. Sometimes it is difficult to identify the mechanism of vibration excitation in complex hydromechanical systems with many sources of oscillation. Simple frequency and correlation analysis does not provide all needed information. It is suggested to perform both usual vibration analysis in frequency domain and check out fluid-induced vibrations and conditions for auto-excitation caused by interaction of reciprocating mechanical parts and pulsing fluid. A case with an ammoniac compressor station is considered as an implication of the method. During the compressor station exploitation vibration has been observed with a frequency of 18.8 Hz that propagates through the system and exceeds the permitted level by 2.23 times. The frequency does not correspond to any known sources of oscillation for the system. The analysis showed that the back valve and attached pipes represent an auto oscillation contour that generates both low-frequency pulsation of fluid and mechanical vibration. It represented results of operational development of the compressor station, using the proposed technique that allowed the reduction of essential vibration of the system and generated noise.

11:40

2aNSb12. The mufflers for exhaust valves of industrial equipment. Valeri V. Lenchine (Inst. of Acoust. of Machines, Samara St. Aerosp. Univ., Moskovskoe shosse 34, Samara 443086, Russia)

The operating of exhaust valves of pneumatic cylinders and other pneumo-automatics causes impulse noise generation. Staff, working on half-automated conveyers and package lines, are often exposed to the high-intensity impulse noise, which is a reason for hearing loss. The design of a silencer for steady flow is well explored but there is not as much attention for an impulse noise attenuator. Moreover, mufflers for discharge valves of industrial equipment must meet additional requirements because an air exhaust emits broadband sound and an implication of mufflers with essential aerodynamic drag results in an operational disorder for pneumo-automatics. The long lifetime and simple maintenance are also significant for a consumer. The muffler's model takes into account an attenuation performance, calculation of a silencer pressure drop, and prediction of mass and size of the device that allows us to work out multi-criteria optimization tasks. There are considered peculiarities of impulse noise muffler design for the packaging pneumatic automates. The mufflers have low aerodynamic drag, are easy to install, inexpensive, and reliable. They provide the best acoustical performance for frequencies 500–4000 Hz. The comparison of effectiveness is represented for the different silencers. There are given examples of successful implementation of the mufflers in industry.

[Note: A Power Plant Acoustics Roundtable will be held on Tuesday afternoon in Session 2pNSd.]

TUESDAY MORNING, 5 DECEMBER 2000

BRIGANTINE/TRIMARAN ROOMS, 7:25 A.M. TO 12:00 NOON

Session 2aNSc

Noise, NOISE-CON and Architectural Acoustics: International Noise Standards

Paul D. Schomer, Chair

U.S. Army Construction Engineering Research Laboratory, P.O. Box 9005, Champaign, Illinois 61826-9005

Chair's Introduction—7:25

Invited Papers

7:30

2aNSc1. The Acoustical Society of America program in standards. Paul Schomer (USA-CERL-CNN, P.O. Box 9005, Champaign, IL 61826-9005, schomer@uiuc.edu)

The Acoustical Society of America (ASA) Standards Secretariat has the delegated authority of the American National Standards Institute (ANSI) to manage the development of national standards, to provide for the USA input to the development of international standards, and to manage the development of certain international standards. First, the ASA Standards Secretariat manages the four ''S'' committees that develop ANSI Standards (national standards) using the consensus method. The four S-committees include S-1 (acoustics), S-2 (mechanical vibration and shock), S-3 (bioacoustics), and S-12 (noise). Second, the ASA Standards Secretariat manages the US representation to several International Organization for Standardization (ISO) and International Electro-technical Commission (IEC) technical committees including ISO TC 108 and 6 of its subcommittees, ISO TC-43 and TC-43-subcommittee 1, and IEC TC 29. Liaison representation from the Society of Automotive Engineers (SAE) and the American Society for Testing and Materials (ASTM) is relied on, in part, to perform this function. Third, the ASA Standards Secretariat provides the international secretariat for ISO TC-108 and two of its subcommittees. This paper discusses the role and program of the ASA Standards Secretariat.

7:50

2aNSc2. Work programs and interrelations of IEC/TC29, ISO/TC43 and ISO/TC43/SC1. Klaus Brinkmann (Physikalisch-Technische Bundesanstalt Braunschweig, Bundesallee 100, 38116 Braunschweig, Germany, klaus.brinkmann@ptb.de)

IEC/TC29 ''Electroacoustics'' and ISO/TC43 ''Acoustics,'' together with its Subcommittee ISO/TC43/SC1 ''Noise,'' each have very distinctive scopes. IEC/TC29 is mainly responsible for specification of performance characteristics of electroacoustical measuring instruments such as microphones, sound-level meters, sound calibrators, and audiometers, while ISO/TC43 and its SC1 primarily deal with measuring methods in audiology and noise control. Close links are, however, established by basically the same procedure for developing standards used in IEC and ISO, by common interest and mutual stimulation with respect to the instrumentation used for acoustical measurements and by overlapping involvement of experts in all three committees. The program of work of each committee will be briefly reviewed. Most actual information will be provided on IEC/TC29, which will have had its meeting just the week before. ISO/TC43 and ISO/TC43/SC1 will meet the following week, thus only allowing a forecast on expected achievements.

8:30

2aNSc3. Current trends in the international standardization of mechanical vibration and shock. Bruce E. Douglas (ISO Tech. Committee on Mech. Vib. and Shock and CSO, Resonance Technologies, 3731 Parke Dr., Edgewater, MD 21037)

Current concerns on world trade (e.g., fair business practices, product quality, and reliability), the environment (e.g., the propagation and radiation of noise from structures), and public safety (e.g., the effects of shock and vibration on humans, civil structures, and vehicles) have placed increased international attention on the development of quality standards in the field of mechanical vibration and shock. Recognizing its importance, the International Organization for Standardization (ISO) has set up a Technical Committee to develop standards in the field of mechanical vibration, shock, and the condition monitoring of machines and structures, ISO/TC 108. Its scope includes the areas of mechanical vibration and shock pertaining to: terminology and nomenclature; actuators, sensors, and associated instrumentation; reduction and control methods; measurement and evaluation of the exposure of humans, stationary structures, vehicles, and machines. In addition, standardization in the field of vibration and shock for data processing, data acquisition, diagnostic measurement methods, transducer calibration, and condition monitoring are actively being developed. This presentation addresses topics on the broad issues of standardization in the field of mechanical vibration and shock with emphasis on (1) societal objectives, (2) economic stakes of diverse constituencies, and (3) structure, process, and future direction of ISO/TC 108.

8:50

2aNSc4. Noise control and the machinery safety standardization program of CEN (European Standardization Committee). Jean R. Jacques (INRS, 30 rue Olivier Noyer, 75680 Paris, cedex 14, France), Roger F. Higginson (Higginson Acoust. Ltd., Bracknell, Berkshire RG129JL, UK), and Patrick Kurtz (BAuA, 44149 Dortmund, Germany)

One major aspect of the European internal market is the free circulation of products with a high level of safety. The European directive on machinery safety includes requirements to design low-noise machinery and to give out quantitative information on noise emission. One way for a manufacturer to get presumption of conformity to the directive is to use so-called harmonized European standards. In order to be harmonized and regarding noise, standards in the CEN machinery program (about 800 work items, each of them corresponding to a particular machinery family) must contain clauses on noise control and noise declaration, and a noise test code based on EN ISO 12001. Noise test codes specify the basic standards to be used for determining the emission sound pressure levels at workstations (among the EN ISO 11200 series) and the sound power level (among the EN ISO 3740 and 9614 series) and for declaring noise emission values (EN ISO 4871), together with the position of workstation(s) and the mounting and operating conditions of a machine during noise measurement. As part of a quality control system, noise consultants are employed by CEN to assess the noise contents of drafts and to assist Working Groups with noise matters.

9:10

2aNSc5. Building acoustic ISO standards and European legal metrology. Hans Goydke (PTB, Bundesallee 1000, 38114 Braunschweig, Germany)

Facilitating trade in the European Union through technical harmonization of products is the purpose of the Construction Product Directive of the European Commission. An appropiate tool to be expected is the development of harmonized international standards which have to be published as national standards in every member state. The cooperation of ISO and CEN succeeded in fulfilling the task, meanwhile, in the field of building acoustics with only a few exceptions. But, before these standards are to be incorporated into the legislation in the countries, for instance regarding legal requirements on the protection against noise in buildings, it has to be verified that the application of these standards shows sufficiently corresponding results when a test object is tested in laboratories in different countries. The results of such tests available, meanwhile, are partly unsatisfying. In order to solve the problems, the development of normative additions to some of the standards, for instance in order to specify more strictly mounting conditions of test objects, are most important work items of actual building acoustics standardization.

9:30

2aNSc6. ASTM Committee E-33 and its role in acoustical and noise standardization. Steven M. Brown (Steelcase, Inc., CD-5E-16, 6100 East Paris S.E., Caledonia, MI 49316, sbrown@scrbd.com)

ASTM Committee E-33 on Environmental Acoustics develops standards on the characteristics and performance of materials, products, systems, and services relating to the acoustical environment and the promotion of related knowledge. The Committee holds jurisdiction for over 50 standards on sound absorption, open plan spaces, sound transmission, the application of acoustical materials and systems, acoustical research, acoustical terminology, mechanical and electrical system noise, and community noise. Among these are the principal building acoustics standards used in North America and other areas. The Committee also provides the U.S. Technical Advisory Group for ISO/TC 43/SC 2 on Building Acoustics. This talk describes and explains E-33's activities and its relationship with other standards development organizations.

9:50

2aNSc7. SAE transportation noise activities. Richard F. Schumacher (GM Proving Ground, 3300 GM Rd., MC 483-356-000, Milford, MI 48380-3726) and Paul R. Donavan (GM Proving Ground, Milford, MI 48380-3726)

The SAE Vehicle Sound Level Form Committee continues to ensure the SAE road vehicle standards are maintained. Currently, there is a trend to move to internationally recognized standards that are technically valid procedures for rating the various products or subsystems. Although SAE has not initiated a development program, the light-vehicle standards group is actively reviewing an ISO project to develop more representatives of urban driving operations. All proposals are being evaluated to ensure that the procedure accommodates all vehicle designs including the US-style electronic-controlled automatic transmissions. An SAE horn and an accessory component committee are both actively involved in updating or developing new standards. SAE, while not involved in any major new initiatives, provides a significant role in review of world procedures. The expansion of our industrial world boundaries requires a continuing adjustment of our priorities to ensure we maintain standards that are relevant and provide the US-based industry an equal opportunity to be active participants in the global markets.

10:10–10:30 Break

Contributed Papers

10:30

2aNSc8. Standards in acoustics of the American Society of Heating, Refrigeration, and Air-Conditioning Engineers (ASHRAE). Richard J. Peppin (Scantek, Inc., 916 Gist Ave., Silver Spring, MD 20910, peppinr@asme.org)

This paper presents the extent of ASHRAE's involvement in sound and noise standards. ASHRAE has only a limited number of formal standards in acoustics but its two volumes of the four-volume set of handbooks contain many items that have become *de facto* standards. The ''Fundamentals Handbook'' and the ''Applications Handbook'' provide tables that show the power loss due to end reflection, attenuation of lined ducts, room criteria approaches and metrics, and other material properties that are used even though, and in spite of the fact that, they never were evaluated as a standard by Canvass or other means. I will summarize the topics in the ASHRAE publications and discuss the limitations and current controversial topics.

10:45

2aNSc9. Environmental noise: From guidelines to standards. Dietrich Schwela (World Health Organization, Occupational and Environ. Health, 20 Ave. Appia, CH 1211 Geneva 27, Switzerland)

In view of the new WHO guidelines for Community Noise, a concept is proposed on deriving environmental noise standards on the basis of response-exposure relationships. WHO's guidelines have set definite guideline values for specific health effects and for specific environments. The WHO guidelines emphasize that exposure-response relationships between politically relevant variables (such as noise-induced social handicap, reduced productivity, decreased performance, absenteeism, drug use, and accidents) could definitely help decision makers to set sensible standards. In the standard-setting procedure, five different risks occur as part of any decision on the value of a standard. Three of these arise from the generalization of statements on susceptible sub-populations to be protected, on adverse effects of noise, and on boundary conditions of epidemiological studies. These risks cannot be quantified. One of the remaining two quantifiable risks is that of wrong measurements, which results from the quality of measurement techniques. The fifth risk relates to the acceptable probability of the occurrence of effects due to noise. With respect to the latter, governments have to fix an acceptable risk value in the standards-setting process. The consequences of combining these risks to a total risk are discussed.

11:00

2aNSc10. Noise impact from motor racing events. Alberto de Leo, Franco Giuliani, and Concetta Fabozzi (Natl. Agency for Environ. Protection of Italy (ANPA), Via V. Brancati 48, 00144 Rome, Italy)

The new Italian rules about noise pollution (General Law no. 447/95 and subsequent decrees) have caused the necessity of carrying out some measurement campaigns in order to observe the acoustic effects of various events and their interaction with territory and population. One of the research studies carried out by ANPA (Italian Agency for Environmental Protection) has dealt with a cognitive survey in motor racing sites and the noise impact that such infrastructures have on the surrounding environment. Measures have been done in and around motordromes, of which two are used for $F1$ races and one for motorcycles races; they differ from one another for territory orography, wind conditions and outskirts urbanization.

11:15

2aNSc11. Problems of evaluation of long-term levels for Corona noise. Tadeusz Wszolek (Univ. of Mining and Metallurgy, Al. Mickiewicza 30, 30-059 Krakow, Poland)

The ac power lines are sources of audible noise (AN), emitted as a result of Corona effect, the intensity of which depends mainly on the weather conditions. Emitted AN mostly varies between 35 and 55 dB(A) at 30-m distance. Corona AN may be considered to be composed of two major components: the low-frequency pure tones and the broadband noise. Whereas the broadband component is more or less stable, the tonal components are subject to great fluctuations. Therefore, their measurement presents certain difficulties. On the other hand, the tonal components are, according to the ISO 1996, accounted for in the AN evaluation by adding a correction to the measured level. Because of seasonal variation of the weather, long-term AN level values are taken for the evaluation, according to the ISO 1996. There is a problem concerning the evaluation of the actual periods of fair and bad weather conditions and their influence on the measurement result. The problem is also not solved by introducing the day/night levels (IEEE Std. 656). Another problem regards the environmental interference. In the paper the effect of possible errors, connected with the above-mentioned problems, on the final results of the long-term level calculations is analyzed.

11:30–12:00

Panel Discussion

TUESDAY MORNING, 5 DECEMBER 2000 PACIFIC SALON B, 8:10 A.M. TO 12:00 NOON

Session 2aSA

Structural Acoustics and Vibration, Signal Processing in Acoustics and Physical Acoustics: Diagnostics of Vibration and Noise in Structures I

Sean F. Wu, Chair
Department of Mechanical Engineering, Wayne State University, 505 Anthony Wayne Drive, Detroit, Michigan 48202

Chair's Introduction—8:10

Invited Papers

8:15

2aSA1. Regularization of inverse problems with evanescent waves. Earl G. Williams (Naval Res. Lab., Code 7137, Washington, DC 20375)

Due to the presence of evanescent waves, acoustic and electromagnetic radiation problems form a class of ill-posed inverse problems when the vibration (electric currents) is to be determined solely from a near-field measurement of the pressure (electric field), as is done in near-field acoustical holography. The problem is ill posed since small changes in the measured pressure lead to large changes in the reconstructed velocity. Regularization of the problem is crucial. We will discuss various regularization methods which have been tested against experimental data and show how they can be modified and tuned in an effort to incorporate the physics of the decaying fields into them. In this way, we hope to identify a single technique which provides near optimum inversion of the ill-posed problem. [Work supported by ONR.]

8:45

2aSA2. Reconstruction of acoustic radiation from an arbitrarily shaped object using the CHELS method. Sean F. Wu and Xiang Zhao (Dept. of Mech. Eng., Wayne State Univ., 5050 Anthony Wayne Dr., Detroit, MI 48202)

The combined Helmholtz equation–least-squares (CHELS) method developed previously by the authors [J. Acoust. Soc. Am. **107**, 2799 (2000)] is used to reconstruct the acoustic field radiated from an arbitrarily shaped object. The main advantage of the HELS method [Wang and Wu, J. Acoust. Soc. Am. **102**, 2020–2032 (1997); Wu, *ibid.* **107**, 2511–2522 (2000)] is its simplicity in expressing the radiated acoustic pressure in terms of an expansion of certain basis functions, whose coefficients are determined by matching the assumed-form solution to the measured acoustic pressures. The number of expansion functions determines that of the measurements, which is small for structures that can be described by spheroidal coordinates. Combining the HELS method with the Helmholtz integral formulation will further allow for reconstruction of acoustic radiation from an arbitrarily shaped structure. This CHELS method is applied to a four-cylinder engine block. The normal component of surface velocity is reconstructed using singular value decomposition and a regularization technique. The corresponding surface acoustic pressure is obtained by solving the Helmholtz integral equation. Once these surface quantities are specified, the normal component of time-averaged acoustic intensity can be determined. Results of this study demonstrate that the CHELS method can become potentially a powerful noise and vibration diagnostic tool for practicing engineers. [Work supported by NSF.]

9:15

2aSA3. Analyzing structure radiation with eigenfunction expansions. J. D. Maynard (Dept. of Phys., The Pennsylvania State Univ., 104 Davey Lab., University Park, PA 16802)

A first step in noise control is to locate the regions on a structure which strongly couple sound energy into the surrounding medium. A powerful tool for accomplishing this is near-field acoustic holography, which can produce a map of the acoustic energy flow surrounding a structure. In the most manageable form of holography, reconstructions of the surface of the source can be obtained if the surface coincides with a level surface of a separable coordinate system, a severe limitation. For treating sources with arbitrarily shaped surfaces a new technique, using spherical wave function expansions and singular value decomposition, was developed at Penn State University. This technique avoids the evanescent waves of plane wave expansions. Computer implementations have become feasible with the availability of high-speed computers. A successful application of the new technique was an explanation of the measured radiation damping of elastic solids vibrating in a high-density gas. [Research was supported by National Science Foundation DMR 9801844 and the US Office of Naval Research.]

9:45

2aSA4. Optimal regularization and resolution limits for the reconstruction of acoustic source strength by inverse methods. Philip A. Nelson and Youngtae Kim (ISVR, Univ. of Southampton, Highfield, Southampton SO17 1BJ, England)

Generalized cross validation (GCV) provides an effective method for the determination of optimal regularization parameters in acoustical inverse problems. The problem of reconstructing acoustic source distributions from field measurements is very often ill-posed.The use of Tikonhov regularization, for example, often suppresses the effect of small singular values in the Green function matrix to be inverted and these are in turn often associated with ''high spatial frequencies'' of the source distribution. The net effect is to produce a useful estimate of the acoustic source strength distribution but with a limited spatial resolution. This paper will explore the relationship between estimation accuracy, spatial resolution, noise level, and source sensor geometry when a range of inverse sound radiation problems is regularized using GCV. [Youngtae Kim is supported by a British Council Grant which is gratefully acknowledged.]

10:15–10:30 Break

Contributed Papers

10:30

2aSA5. An outline of a proposed acoustic holography system based on the Helmholtz equation–least-squares (HELS) method. Nassif E. Rayess and Sean F. Wu (Dept. of Mech. Eng., Wayne State Univ., Detroit, MI 48202)

The Helmholtz equation–least-squares (HELS) method [Wang and Wu, J. Acoust. Soc. Am. **102**, 2020–2032 (1997); Wu, J. Acoust. Soc. Am. **107**, 2511–2522 (2000)] has been proven a reliable acoustic holography method. The main advantages are its simplicity in formulations, flexibility in measurement locations, and uniformity in algorithm for a variety of source geometry and conditions. This paper presents a complete acoustic holography system based on the HELS method that makes full use of these advantages. This system has been under development as part of an ongoing research at the Acoustics, Vibration, and Noise Control Laboratory at Wayne State University. The physical setup will be described which includes the front-end multichannel data acquisition system and the associated hardware. Experimental results on reconstruction of the radiated acoustic fields from a complex source using this system will be demonstrated. Comparisons of this system with other acoustic holography systems currently available will be made. [Work supported by NSF.]

10:45

2aSA6. Why can spherical harmonics be used to describe the sound field generated by nonspherical sources? Nassif E. Rayess and Sean F. Wu (Dept. of Mech. Eng., Wayne State Univ., Detroit, MI 48202)

The spherical harmonics in conjunction with the spherical Hankel functions form a complete set of spherical basis functions, which can fully describe the radiated sound field from a vibrating object. The vibrating object itself need not be spherical as long as the exterior region is beyond a minimum sphere that completely encloses the object. Such expansion solutions are not supposed to be valid inside the minimum sphere. However, experimental results [Rayess and Wu, J. Acoust. Soc. Am. **107**] using the Helmholtz equation–least-squares (HELS) method [Wang and Wu, J. Acoust. Soc. Am. **102**] have shown that a modified set of spherical basis functions can accurately describe the sound field well inside the minimum sphere. That modified set is obtained by Gram–Schmidt orthonormalization of the spherical basis functions with respect to the source surface. It is shown that the velocity fields corresponding to these modified basis functions form an orthogonal and complete set, which can uniquely describe the velocity field on the surface. The proof given in this paper

will provide a solid foundation for the establishment of the HELS method as a potentially viable noise diagnostic tool. [Work sponsored by NSF and Ford Motor Company.]

11:00

2aSA7. Determination of the optimal number of expansion terms in the HELS method. Byoung-Duk Lim (Dept. of Mech. Eng., Yeungnam Univ., 214-1 Dae-dong, Gyongsan, Kyungbuk, 712-749, Korea) and Sean F. Wu (Wayne State Univ., Detroit, MI 48202)

The Helmholtz equation least-square (HELS) method has been proven to be an efficient way of reconstructing acoustic radiation from a vibrating structure [Wu, J. Acoust. Soc. Am. **107**, 2511–2522 (2000)]. In this method, the optimal number of expansion terms is determined by a heuristic method of searching the minimum–norm errors of the reconstructed acoustic pressures with respect to the measured data. While this approach works, it is time consuming. In this paper, a new method of *a priori* determination of the optimal number of expansions is developed based on the eigenvalue analysis of the measured information. Assuming a Gaussian white measurement noise, the hypothesis of equal noise eigenvalues is tested based on the statistical properties of the eigenvalues. This hypothesis test yields an upper bound of the optimal number of expansions in general. Also developed is a method for testing the correlation coefficients after eliminating some principal components, which gives a lower bound of the optimum number of expansions. Since the curve of minimum–norm errors in the previous work shows local minima when the number of terms is squares of the integer, the range produced by the present method leads to only a few candidates for the optimal number of expansions. Examples of reconstructing the radiated acoustic pressure field from a vehicle front end using the optimal number of expansion terms thus obtained are demonstrated in parallel with the results of the *a posteriori* method discussed by Wu. [Work supported by NSF.]

11:15

2aSA8. Vibro-acoustography of small spheres. Shigao Chen, Mostafa Fatemi, and James F. Greenleaf (Dept. of Physiol. and Biophys., Mayo Clinic, Rochester, MN 55905, chen.shigao@mayo.edu)

Vibro-acoustography is a method for imaging the acoustic energy emitted from objects in response to an oscillatory radiation force produced by two interfering focused beams of ultrasound [M. Fatemi and J. F. Greenleaf, Science **280**, 82–85 (1998)]. To facilitate quantitative study, a model is presented that describes the behavior of an elastic sphere in two plane ultrasound waves of slightly different frequencies. The vibrating velocity and the resulting acoustic emission of the sphere at the difference frequency are derived from the radiation force and the impedance of the sphere. For small steel spheres in water, the velocity and emission are predicted to have a simple relationship with the difference frequency and the size of the sphere. Experiments on small stainless steel spheres of different size are carried out for a number of difference frequencies, with a laser interferometer to measure the velocity of the sphere and a calibrated hydrophone to measure the acoustic emission from the sphere. The experimental results fit the model description well. This means that vibro-acoustography may be useful in quantitative measurement.

11:30

2aSA9. Backscattering enhancements associated with antisymmetric Lamb waves on circular plates in water: Observations and imaging of the plate surface using acoustic holography. Brian T. Hefner and Philip L. Marston (Phys. Dept., Washington State Univ., Pullman, WA 99164, bhefner@mail.wsu.edu)

When a tilted circular plate is illuminated with high-frequency sound in water, significant enhancements occur when the angle of incidence corresponds to a Lamb wave coupling angle. Previously, scattering associated with the lowest-order symmetric Lamb wave, s_0, was studied using acoustic holography and modeled using quantitative ray methods [B. T. Hefner and P. L. Marston, J. Acoust. Soc. Am. **107**, 2847 (2000)]. The present investigation focuses on backscattering associated with the antisymmetric Lamb waves. Above the cutoff frequency of the first-order antisymmetric wave, a_1, enhancements are observed which are associated with both the a_1 wave and the a_0 wave. For these enhancements, however, mode conversion between these modes plays a significant role. To examine the mode conversion and to identify the scattering mechanisms, acoustic holography was used to image the plate surface when the plate is ensonified at the a_1 wave coupling angle. From the imaged wave field, the enhancement at this angle involves a wave which moves around the circumference of the plate: a ''whispering gallery'' mode involving the a_1 wave or a flexural edge wave. Both of these possibilities are considered as well as a possible scattering mechanism at the a_0 wave coupling angle. [Work supported by the Office of Naval Research.]

11:45

2aSA10. Diagnostics of vibration and noise via Fourier transforms: A new approach. Leonid M. Gelman (Dept. of Nondestructive Testing, Natl. Tech. Univ. of Ukraine, 37, Peremogy Pr., Kiev, 252056, Ukraine), Ivan V. Petrunin (Natl. Tech. Univ. of Ukraine, Kiev, 252056, Ukraine), and Vladimir T. Shirkov (''Motor-Sich'' Co., Zaporozh'e, 01202, Ukraine)

A new general optimal approach and new diagnostic features are proposed, for those cases where one- and multidimensional Fourier transforms are used for diagnostics of vibration and noise. For the forced oscillation vibroacoustical diagnostics method with Gaussian excitation, the optimal nonlinear transformation of the proposed features is received. It was shown (1) The power spectral density is optimal transformation for considered diagnostics only for specific situations; (2) The phase spectrum is not optimal transformation for considered diagnostics; and (3) The proposed approach provides the increment of diagnostics effectiveness in contrast with usage of power spectral density. This is in contrast to most applications concerning diagnostics of vibration and noise, where power spectral density is used.

Session 2aSC

Speech Communication: Articulation (Poster Session)

Kristin K. Tjaden, Chair

Department of Communicative Disorders and Sciences, The State University of New York at Buffalo, 122 Cary Hall, 3435 Main Street, Buffalo, New York 14214-3005

Contributed Papers

All posters will be on display from 9:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 9:00 a.m. to 10:30 a.m. and contributors of even-numbered papers will be at their posters from 10:30 a.m. to 12:00 noon.

2aSC1. A cross-linguistic acoustic study of fricatives. Matthew K. Gordon, Paul Barthmaier, and Kathy Sands (Dept. of Linguist., Univ. of California, Santa Barbara, Santa Barbara, CA 93106)

This work presents results of an acoustic study of fricatives in 7 languages (Aleut, Chickasaw, Hupa, Montana Salish, Scottish Gaelic, Toda, and Western Apache), all of which contrast fricatives made at several places of articulation. Measurements of the frequency of spectral peaks and centroid frequencies indicate many similarities between the languages in the acoustic properties defining the fricatives. Some of the principal findings are the following. Alveolar sibilants typically have the highest spectral peak and centroid frequency. Lateral and palatoalveolar fricatives have spectral peaks and centroids intermediate in frequency between alveolar sibilants and backer fricatives. Among the back fricatives, peaks and centroids of uvulars are characteristically lower than those of velars. Rounding of back fricatives induces further lowering of peaks and/or centroids. Contrasts in backness and rounding among the back fricatives are also associated with differences in $F2$ of the following vowels: $F2$ values are lower following uvulars than velars, and lower following rounded than unrounded fricatives. Labiodental fricatives typically have flat spectra with poorly defined spectral peaks. Finally, the contrast between lateral fricatives and palatoalveolar sibilants is variably realized, depending on language and speaker, as a difference in the location of spectral peaks and/or centroid frequency.

2aSC2. The noise spectrum characteristics of retroflexed consonants in Mandarin Chinese. Jing-Yi Jeng (Dept. of Special Education, Natl. Tainan Teacher College, Taiwan, ROC)

Retroflexed consonants are the most complex and difficult sounds for Mandarin speakers. The four retroflexed consonants in Mandarin are produced with the tip of the tongue retracted and curled up close to the positions of the hard palate. The purpose of this study was to investigate the acoustic characteristics of retroflexed consonants for normal speakers. The moment analysis and the noise duration of retroflexed and nonretroflexed fricatives and affricates were studied for 10 adult speakers. Moment analysis is a statistical approach for the power spectrum. The results showed that retroflexed consonants had lower mean noise frequencies (M1) than nonretroflexed consonants. The standard deviation of noise energy frequencies (M2) of retroflexed consonants was smaller than those of nonretroflexed consonants. Results will be discussed in terms of the linkage between acoustic data and tongue positions. The discriminate analysis by using moment and duration parameters to classify word-initial retroflexed consonants was also reported.

2aSC3. New effects of laryngeal configurations on $f0$: Voiceless stops in Korean. Sunyoung Oh (Dept. of Linguist., Univ. of British Columbia, E270-1866 Main Mall, Vancouver, BC V6T 2E1, Canada, sunyoh@interchange.ubc.ca)

Korean has three sets of voiceless stops in word-initial position which are distinctive from one another in terms of laryngeal adjustments: lenis unaspirated, fortis unaspirated, and aspirated [P. Ladefoged and I. Maddiedson, 47–101 (1996)]. It has been observed that, in Standard Korean, the aspirated stops have higher $f0$ value than the fortis stops, while both stops have higher $f0$ value than the lenis stops [D. Silva, 11–34 (1998)]. It could be argued that this indicates a universal correspondence between laryngeal configurations for voicing and pitch—at least within Korean. However, the current paper reveals a different observation. In South Kyengsang dialect in Korean, little difference is found in the $f0$ value between the aspirated and the fortis stops, while the $f0$ of the lenis stops is relatively lower, similar to Silva's (1998) finding. A total of 600 tokens, by five speakers, were recorded and digitized, then analyzed using Macquirer software for pitch tracking. The results show that the $f0$ following the fortis stops are consistently higher than that of Standard Korean, regardless of the place of articulation, and the higher $f0$ of the fortis stops patterns with the lower $f0$ of the lenis stops among speakers.

2aSC4. Acoustic evidence of a phonetics-phonology mismatch in Nuu-chah-nulth. Ian L. Wilson (Dept. of Linguist., Univ. of British Columbia, Buchanan E270-1866 Main Mall, Vancouver, BC V6T 1Z1, Canada)

Measurements were made of the acoustic properties of the pharyngeal ʕ in Ahousaht Nuu-chah-nulth (Nootka), an endangered language spoken on Vancouver Island. One goal of the study was to test whether phonetics and phonology make the same predictions regarding the classification of ʕ. Although the phonological data largely favors an analysis of ʕ as a type of glottalized pharyngeal stop, the results of acoustic measurements suggested that it is a glottalized glide. These acoustic measurements included a comparison of the waveforms and spectrograms of ʕ to those of glottalized resonants and glottalized stops. Three types of acoustic evidence were found to support the claim that ʕ is a glottalized glide. Specifically, the timing of glottalization in ʕ was like that of a glottalized resonant, rather than a glottalized stop; in several tokens full stop closure was not achieved; and the duration of the formant transition of ʕ is more like that of a glide than a stop. This acoustic evidence helps to illustrate one example of a phonetics-phonology mismatch, and the fact that there is a mismatch supports a modular view of grammatical organization. [Work supported by SSHRCC and NSERC Grants.]

2aSC5. Cues for place and palatalization in Russian stops. Colleen Richey (Dept. of Linguist., Stanford Univ., Margaret Jacks Hall Bldg. 460, Stanford, CA 94305-2150)

Using acoustic measurements and native speaker perceptions, the cues for place of articulation of Russian plain and palatalized stops were investigated. The acoustic measurements were made of formant transitions, burst duration, and spectral burst shape. They showed that formant transitions distinguish place of articulation of plain consonants, but not palatalized consonants. Differences distinguishing place of articulation were found in the burst durations and in the ratio of energy between the upper and lower portions of the burst spectra. In the perception experiment, native Russian speakers listened to various consonant–vowel stimuli. Hearing only the formant transitions, listeners were able to identify place of articulation of just the plain stops. When they heard stimuli with conflicting cues for place in the burst and formant transitions, in the majority of the cases, they identified plain consonants based on the formant transitions, but they identified palatalized consonants based on the burst. It is concluded that while formant transitions do provide the strongest cues for place among plain stops, they do not provide cues for place among palatalized stops. The cues for place of articulation of palatalized consonants are contained completely in the duration and spectral shape of the burst.

2aSC6. Dialect differences in the temporal characteristics of vowels: A comparison of standard (Parisian) and Swiss French. Joanne L. Miller, Michèle Mondini (Dept. of Psych., 125 NI, Northeastern Univ., Boston, MA 02115, jlmiller@neu.edu), François Grosjean (Univ. of Neuchâtel, Switzerland, francois.grosjean@lettres.unine.ch), and Jean-Yves Dommergues (Univ. of Paris, 7, France)

Earlier we reported a dialect difference in the use of temporal information for vowel perception: Native speakers of Swiss French used temporal as well as spectral information when identifying /o/ versus /ɔ/, whereas native speakers of standard (Parisian) French used only spectral information [J. L. Miller and F. Grosjean, Language Speech **40**, 277–288 (1997)]. We interpreted this dialect difference in terms of the more prominent role that vowel duration plays overall in the phonological system of Swiss French compared to standard French. To investigate further the basis of the dialect effect, we have been measuring the duration of /o/ and /ɔ/ in monosyllabic words for native speakers of the two dialects. Our findings to date indicate a robust dialect effect in production: The duration difference between /o/ and /ɔ/ is substantially larger and more consistent in Swiss French than in standard French. Thus the perceptual dialect effect for /o/ and /ɔ/ we reported earlier reflects both a specific difference in the temporal characteristics of this vowel pair and an overall difference in the role of vowel duration in the phonological systems of the two dialects. [Work supported by NIH/NIDCD, Swiss NSF.]

2aSC7. Gemination in Italian: the GEMMA project. Maria-Gabriella Di Benedetto (Univ. of Rome La Sapienza, Infocom Dept., via Eudossiana 18, 00184 Rome, Italy, gaby@acts.ing.uniroma1.it)

The Gemination project (GEMMA) started at the University of Rome La Sapienza, Italy, in 1992, examined gemination in Italian. The analyzed consonants were stops, liquids, fricatives, nasals, and affricates. GEMMA was completed in 2000. Partial results were previously reported [stops: Esposito, DiBenedetto (1999), liquids: Argiolas *et al.* (1995), fricatives: Giovanardi, DiBenedetto (1998), nasals: Mattei, DiBenedetto (2000), affricates: Faluschi, DiBenedetto (2000)]. The major findings were coherent in all studies: gemination is revealed by time-related parameters, namely, the consonant lengthening and the preconsonant vowel shortening in geminate forms. Relevance of acoustic correlates was confirmed by perceptual experiments [A. Esposito and M. G. DiBenedetto, J. Acoust. Soc. Am. **106**, 411–422 (1999); F. Argiolas *et al.*, J. Acoust. Soc. Am. **97**, 3418 (1995); M. Giovanardi and M. G. DiBenedetto, Eur. J. Language Speech 1–13 (1998), http://web-sls.essex.ac.uk/web-sls; M. Mattei and M. G. DiBenedetto, Eur. J. Language Speech 1–14 (2000), http://web-sls.essex.ac.uk/web-sls; S. Faluschi and M. G. DiBenedetto, Eur. J. Language Speech (submitted)].

2aSC8. Effects of speaking rate on vowel formant trajectories. Jeff Berry and Gary Weismer (Dept. of Communicative Disord. & Waisman Ctr., Univ. of Wisconsin—Madison, Goodnight Hall, Madison, WI 53706, weismer@waisman.wisc.edu)

The effects of speaking rate and vowel context on formant frequencies have traditionally been assessed by analyzing the variation in some operationally defined ''target'' frequency value. The use of a single frequency value to describe the behavior of a formant has proven useful for distinguishing among vocalic nuclei and identifying ''target undershoot'' effects of speaking rate. We have, however, only limited understanding of the characteristic effects of speaking rate change on formant trajectories. In our work, analysis of more complete representations of the time-varying changes in formant frequencies throughout vocalic nuclei have revealed patterns of formant change across speaking rates that differ qualitatively and quantitatively between speakers. These data suggest speaker-specific strategies for accomplishing speaking rate change. We have also observed that the acoustic effects of rate change within speakers differ qualitatively and quantitatively between different ranges of the vowel duration continuum. Thus, the acoustic effects of small changes in speaking rate at speaking rate extremes are not comparable to quantitatively identical changes at habitual rates. Since changes in speaking rate are commonly thought to correspond with changes in the overlap of neighboring articulatory gestures, these findings are pertinent to modeling coarticulation. [Work supported by NIH DC 0149-10.]

2aSC9. Spectral characteristics of /s/ and /ʃ/: Rate effects in Parkinson's disease. Kris Tjaden, Jessica Kleinhaut, and Alexa Kozak (Dept. of Communicative Disord. and Sci., The State Univ. of New York at Buffalo, 3435 Main St., Buffalo, NY 14214-3005)

Comtemporary speech production theory proposes that speakers adjust rate to enhance or maximize acoustic-perceptual distinctiveness for listeners. In fact, the suggestion that rate impacts acoustic-perceptual distinctiveness is one rationale for using rate manipulation as a therapy to improve speech intelligibility in dysarthria. The relationship between rate manipulation, acoustic measures of production, and perceptual impressions of speech in dysarthria is poorly understood, however. For example, some studies report rate-induced acoustic changes in speech that are systematically tied to perceptual impressions. Other studies report no effect of rate on speech production characteristics in dysarthria. Further study of the relationship between speech production measures and rate variation in dysarthria therefore appears warranted. The present study examined spectral characteristics of word-initial /s/ and /ʃ/ produced at a variety of rates by speakers with hypokinetic dysarthria associated with Parkinson's disease. Healthy speakers were studied for comparison purposes. Fricatives were produced in the words *shores*, *shies*, *size*, and *soys* embedded in a carrier phrase. A graded rate task was used to elicit an extensive range of rates. Spectral moment coefficients were used to characterize fricatives; phrase duration was used to index rate. The relationship between rate variation and spectral moment coefficients was quantified via regression analysis.

2aSC10. An evaluation of phonetic working space in normal geriatrics and persons with motor speech disorders. Kate Bunton, John Westbury, and Gary Weismer (Waisman Ctr., Univ. of Wisconsin—Madison, 1500 Highland Ave., Madison, WI 53705, bunton@waisman.wisc.edu)

Speakers with dysarthria seem to have smaller than normal phonetic working spaces. Their vowels exhibit unusual and often reduced formant transitions and slopes, and areas enclosed by their ''vowel locations'' plotted in an ($F1$,$F2$) plane are abnormally small. Presumably, reduced

formant transitions reflect small movements of individual articulators. Reduced lip and jaw movements have been reported for speakers with Parkinson's disease and amyotrophic lateral sclerosis. However, comparable reports for tongue movement are not available. In this study, x-ray microbeam data were used to compare the acoustic and articulatory phonetic working spaces of neurologically normal and dysarthric speakers. Measurements of formant frequencies, and sagittal-plane locations of tongue fleshpoints, were made for the English vowels /i ae a u/ at the temporal midpoint of the vowel nucleus in three different words for each of ten normal and seven dysarthric speakers. Speakers' articulatory phonetic working spaces were defined as areas enclosed by tongue fleshpoint locations. Data obtained from this study provide an enhanced appreciation of the articulatory bases of acoustic features of dysarthric speech. [Work supported by 1-R01-DC03723.]

2aSC11. Observing the link between speech production and speech perception from disordered speech. Emily Q. Wang (Dept. of Commun. Disord. and Sci., Rush Univ., 1653 W. Congress Pkwy., Ste. 203 SENN, Chicago, IL 60612), Yi Xu (Northwestern Univ., Evanston, IL 60208), Katie Kompoliti, and Christopher G. Goetz (Rush Univ., Chicago, IL 60612)

It was recently reported that individuals with idiopathic Parkinsons Disease (PD) and hypokinetic dysarthria were able to signal sentence focus using acoustic parameters such as $f0$, intensity, and vowel duration that normal speakers use. However, it was also found that PD speakers did not show all the $f0$ patterns that normal speakers use. For example, although the $f0$ range of the word under focus was expanded, the $f0$ range of the postfocus word was not always suppressed. In this paper, the $f0$ patterns produced by ten speakers with idiopathic PD (Hoehn and Yahr stage 3 and 4, off) and mild, or mild-moderate, or moderate hypokinetic dysarthrias will be examined. These patterns will also be correlated to their perception by normal speakers. The following issues will be addressed: (1) Can listeners correctly perceive focus produced by the PD speakers when the $f0$ patterns are similar to those used by normal speakers? (2) Can listeners correctly perceive focus when there are missing cues in the $f0$ patterns (e.g., lack of post-focus suppression)? (3) Can misperception result from the speakers attempt to compensate for their articulatory limitations due to the disease process? Implications of the findings will be discussed in terms of speech motor control.

2aSC12. On the correlation between articulatory and acoustic data. Jintao Jiang, Abeer Alwan (Elec. Eng. Dept., Univ. of California, Los Angeles, Los Angeles, CA 90095), Patricia Keating (Univ. of California, Los Angeles, Los Angeles, CA, 90095), Lynne E. Bernstein, and Edward Auer (House Ear Inst., Los Angeles, CA 90057)

In J. Acoust. Soc. Am. **107**, 2904 (2000), we investigated the correlations between external facial movements, tongue movements, and speech acoustics using a database of consonant–vowel syllables spoken by one female and two male talkers. The intelligibility of the talkers, based on visual information, was judged by hearing-impaired individuals. In this study, we extend the analysis to two sentences and a fourth talker (female); each sentence was repeated five times. Two of the talkers (one male and one female) had high-intelligibility ratings while the other two had low ratings. A Qualisys (optical motion capture system) and an EMA system were used to characterize orofacial and tongue movements, respectively. The articulatory and acoustic data streams were recorded simultaneously and aligned. Acoustic features were represented by line spectral pairs. A multilinear regression technique was applied to quantify the correlation between the data. A universal estimator was obtained with training data, using a jack-knife approach, and then applied to the test data. Correlations results were speaker dependent, and, on average, the highest correlations were between the optical and EMA data. Detailed results, including a comparison between the results for CVs and sentences, and implications will be discussed. [Work supported in part by NSF KDI award 9996088.]

2aSC13. Tongue surface dynamics during speech and swallowing. Jordan R. Green, Steven L. Pittelko, and Yutsai Wang (Dept. of Commun. Disord., Univ. of Wisconsin–Madison, 1975 Willow Dr., Madison, WI 53706)

This investigation characterizes tongue surface dynamics that underlie phonemic variation and that distinguish speech from swallowing. Vertical displacements of pellets affixed to the tongue were extracted from the x-ray microbeam database [Westbury, J. X-ray Microbeam Speech Production Database Users Handbook, Version 1 (1994)], which contains articulatory kinematic data from 57 typical speakers. Participants recited 21 vowel–consonant–vowel (VCV) combinations, read the Grandfather Passage, and swallowed 10 cc of water. Consonantal context was manipulated in the VCV utterances as a means to describe the range of tongue dynamics produced during speech. For each task, tongue dynamics was quantified by performing zero-lag cross correlations on selected pellet pairs. A coupling index was then computed by scaling the derived coefficients by a multiplier that reflected the amplitude of each displacement signal. A wide range of movement coupling among tongue pellets was observed across tasks. Phonemic differentiation in vertical tongue dynamics was observed as coupling varied predictably across marker pairs with place of articulation. Moreover, tongue displacements for speech and swallowing clustered into distinct groups based on their coupling profiles. The strengths and weaknesses of the coupling index for characterizing tongue surface dynamics across multiple speakers are considered.

2aSC14. Vocal-tract model experiment for illustrating vowel production. Soo-Ki Hong and S. W. Yoon (Acoust. Res. Lab., Dept. of Phys., Sung Kyun Kwan Univ., Suwon 440-746, Republic of Korea)

Everyone can articulate the same words even though each individual has unique vocal organ. Findings of common features in the same vowels articulated by everyone are very useful for speech recognition and reproduction. In this study a vocal tract is modeled with a cylindrical silicone tube to understand the mechanisms of vowel production. The constrictions on the tube are varied with locations, lengths, and degrees. The resonance frequency shifts of vocal-tract depend on the length ratio and the cross-section area ratio of the constriction in the tube. The resonance frequencies show good correlation with the formants measured in speech. [One of the authors (SWY) was supported by the BK21 project.]

2aSC15. Simulations of voice onset–offset hysteresis in /aha/ utterances. Jorge C. Lucero (Dept. of Mathematics, Univ. of Brasilia, Brasilia DF 70910-900, Brazil, lucero@mat.unb.br) and Laura L. Koenig (New York Univ.)

The subcritical Hopf bifurcation model for voice onset–offset hysteresis [J. C. Lucero, J. Acoust. Soc. Am. **105**, 423–431 (1999)] is investigated considering the glottal area as control parameter. A two-mass model of the vocal folds coupled to a two-tube approximation for the vocal tract is used for producing oral airflow simulations. The glottal rest area is varied in an abduction–adduction pattern to reproduce smoothed oral airflow records measured during /aha/ production [L. L. Koenig and R. S. McGowan, J. Acoust. Soc. Am. **100**, 2689(A) (1996)]. The resultant simulations show good agreement with the unsmoothed measured records, with a clear hysteresis effect: vocal-fold oscillation stops during glottal abduction at a higher value of the glottal area than the value at which it starts during the adduction. The dependence of the onset–offset hysteresis and voice parameters on vocal-fold stiffness (Q factor), subglottal pressure, and male–female laryngeal configurations is analyzed and illustrated with bifurcation diagrams. [Work supported by CNPq, Brazil, and NIH.]

2aSC16. An EPG study of allophonic variation in the gestural magnitude of English lingual consonants. Bryan Gick (Dept. of Linguist., Univ. of British Columbia, E270-1866 Main Mall, Vancouver, BC V6T 1Z1, Canada, gick@interchange.ubc.ca)

Previous work on English glides and liquids has shown that certain (consonant) gestures are most constricted in initial allophones, least constricted in final allophones, and constricted to an intermediate degree in ambisyllabic or resyllabified allophones. It is not known, however, whether this generalization applies to other consonants such as stops and fricatives. As the point-tracking and imaging techniques used in previous studies are not well suited for measuring these types of segments, the present paper will present the results of an EPG study testing the degree of tongue contact during lingual consonants in these three syllable positions. Results of a pilot study using two subjects and a reduced set of consonants were mixed, failing to replicate previous findings for some previously measured consonants (e.g., /n/), while supporting the predicted pattern for others not previously measured (e.g., /z/). Results will be presented for additional subjects using the full inventory of lingual consonants. [Work supported by NSERC.]

2aSC17. A study of compensation for a labial perturbation of the vowel /u/ using an area function model of the vocal tract. Brad H. Story (Dept. of Speech and Hearing Sci., Univ. of Arizona, P.O. Box 210071, Tucson, AZ 85721-0071, bstory@u.arizona.edu)

This study reports on the ability of an area function model of the vocal tract [Story and Titze, J. Phonetics **26**, 223–260 (1998)] to compensate for a labial perturbation. The area function model, which consists of a neutral shape and two weighted orthogonal shaping patterns (modes), is first configured for a typical production of /u/. An effective lip tube based on the labial perturbation experiment reported by Savariaux *et al.* [J. Acoust. Soc. Am. **98**, 2428–2442 (1995)] is then imposed on the area function model (by forcing the lip section to be a constant cross-sectional area). Frequency responses are computed for both the normal and the perturbed shapes. Next, an attempt was made to vary the weighting values for the two shaping patterns of the area function model so that the first two formants were moved back toward their unperturbed frequency locations. The result was that weighting values can be determined that do move the formants in the appropriate direction for compensation. Additionally, the resulting area function shows characteristics similar to those presented by Savariaux *et al.*; that is, a movement of the constriction location posteriorly and a widening of the anterior portion of the front cavity. [Work supported by NIH R01-DC04789-01.]

2aSC18. Electropalatographic and acoustic measures of adaptation to palatal perturbation. Wendi A. Aasland, Shari R. Baum (School of Commun. Sci. & Disord., McGill Univ., 1266 Pine Ave. W., Montreal, QC H3G 1A8, Canada, c3cr@musica.mcgill.ca), and David H. McFarland (Universite de Montreal, Montreal, QC, Canada)

The present investigation examined adaptation to a palatal perturbation in [s] production, as reflected in both electropalatographic (EPG) and acoustic measures. The goal was to explore the development of compensatory motor programs during the production of individual fricative segments, as well as over a 1-h period of focused practice. Thirty repetitions of the syllable [sa] were produced at each of five time intervals (times 0, 15, 30, 45, and 60 min) by a speaker wearing a specially designed pseudopalate with a 6 mm buildup of acrylic at the alveolar ridge. At times 0 and 60, baseline measures of [sa] production with a thin pseudopalate and with no palate in place were also recorded to characterize unperturbed articulation. Acoustic analyses focused on centroid frequencies for the [s] productions, while a wide range of EPG analyses were conducted including measures of groove length and width, location and duration of maximum constriction, and variability in tongue-palate contact patterns throughout the fricative production. Preliminary analyses indicate a high degree of variability throughout the practice period. Results are discussed in relation to the development of adaptive articulatory programs in speech motor control and the articulatory configurations necessary for adequate [s] production. [Work supported by NSERC.]

2aSC19. The use of functional data analysis to study variability in children's speech. Laura L. Koenig (Dept. of Speech-Lang. Pathol. and Audiol., New York Univ. and Haskins Labs) and Jorge C. Lucero (Univ. of Brasilia)

A common finding in studies of children's speech is increased variability relative to adults. This result may be observed well after the age when basic phonological contrasts are established. Researchers debate the sources of this variability, but most agree that the study of variability, and its decrease over time, may provide insight into motor control development. Recent data on oral airflow during /h/ in 5-year-olds and adults revealed that the children presented greater relative amplitude variability than the adults, and seemed to show less regularity in the rates of ad/abduction for /h/. This presentation applies functional data analysis (FDA) to the airflow data from two women and one 5-year-old speaker. In FDA, signals are nonlinearly normalized using a smooth warping function of time, and separate indices are obtained for phasing (temporal) and amplitude variability. Results indicate that the child shows greater variability in both temporal and amplitude measures, but the extent of adult–child differences depends somewhat on the specific methods of normalization and alignment. FDA shows potential as a means of gaining more detailed information about the nature of variability in children's speech data. [Work supported by NIH and CNPq, Brazil.]

Session 2aUW

Underwater Acoustics and Acoustical Oceanography: High-Frequency Sediment Acoustics and Associated Sediment Properties and Processes I

Eric I. Thorsos, Chair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

Chair's Introduction—7:55

Contributed Papers

8:00

2aUW1. SAX99: An experiment in high-frequency sediment acoustics. Eric I. Thorsos (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

In the fall of 1999, ''SAX99'' (for sediment acoustics experiment—1999) was performed at a site 2 km from the shore on the Florida Panhandle in water 18–19 m in depth. The seafloor sediment was medium sand, and acoustic frequencies were mainly in the 10–300-kHz range. A wave-induced ripple field was present at the site. The main goals in SAX99 were to quantify acoustic backscattering from the sediment, acoustic penetration into the sediment (above and below the critical angle of 30 deg), and acoustic propagation within the sediment. Extensive environmental characterizations were made at the experiment site. Quantities that enter into a Biot model description of the sediment were measured. Sediment variability was measured to centimeter scales, giving the most complete coverage for the frequency range of 10–50 kHz. An overview of the entire SAX99 measurement program will be given. The range of acoustic measurements will be briefly described, as well as the scope of measurements of the seafloor physical properties and of biological processes within the sediment and the water column. Initial results of this work will be described in separate papers. [Work supported by ONR.]

8:15

2aUW2. Buried target detection using synthetic aperture sonar. Kerry W. Commander and John E. Piper (Coastal Systems Station, Code R21, 6703 W. Hwy. 98, Panama City, FL 32407-7001, commanderkw@ncsc.navy.mil)

During the Sediment Acoustics Experiment in 1999, the Coastal Systems Station fielded a synthetic aperture sonar (SAS) operating at 20 kHz. The SAS was able to detect and image 1.5×0.5 m cylinders at burial depths up to 50 cm and grazing angles well below the critical angle. In order to clarify the effect of the rough surface (sand ripple field) on sound propagation into the sediment, an experiment was recently conducted using five 10-in. fluid-filled target spheres that were buried at depths ranging from 5 to 50 cm below the top of the spheres. The interior fluid was chosen to maximize the target strength of the spheres (approximately -11 dB) when buried in sediment with a sound speed of 1750 m/s. Using aspect-independent targets enables the determination of the sound level transmitted into the sediment as a function of the angle between the transmitted acoustic beam and the sand ripple direction. Complete 360 deg azimuthal patterns of target echo levels are presented, illustrating the variation due to transmit/receive azimuthal angle with respect to the ripple structure. These azimuthally dependent echo level patterns are presented for several incident grazing angles, all of which are subcritical. [Work supported by ONR Code 321TS.]

8:30

2aUW3. Applied Physics Laboratory participation in SAX99: Experiment design and execution. Kevin L. Williams, Darrell R. Jackson, Dajun Tang, and Eric I. Thorsos (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

The main goals of the APL program in SAX99 were to measure and improve our ability to model acoustic propagation within, high-frequency backscattering from, and penetration into sand sediments. To prepare for these measurements, new equipment and experimental procedures were developed. For the penetration studies, simulations were used extensively to guide the experiment design in order to ensure that the measurements would be useful for addressing our goals. Illustrations will be given of how simulations were used to support the experiment design. The APL experimental equipment used in SAX99 will be described, and the experimental procedures will be presented. Finally, the resulting data set will be summarized. [Work supported by ONR.]

8:45

2aUW4. Acoustic penetration at subcritical grazing angles: Measurements and comparison with simulations based on perturbation theory. Darrell R. Jackson, Kevin L. Williams, Dajun Tang, and Eric I. Thorsos (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

As part of the SAX99 experiment, a buried hydrophone array was deployed together with a movable tower with attached sources covering the frequency range 11–50 kHz. This system was used to examine subcritical penetration into the sediment. For incident grazing angles below the critical angle, scattering dominates the penetrating field. Comparisons with simulations based on perturbation theory show that the penetration is predominately the result of diffraction by the low-amplitude ripple field prevalent at the SAX99 site. Simulations predict a cutoff effect as a function of frequency and grazing angle that is found in the data, and predict changes in penetration as a function of ripple field amplitude that are consistent with those observed. [Work supported by ONR.]

9:00

2aUW5. Refraction and scattering in sandy sediments. Nicholas P. Chotiros, D. Eric Smith, and James N. Piper (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029, chotiros@arlut.utexas.edu)

Signals collected from a buried acoustic receiving array in a sandy sediment off the west coast of Florida, as part of the Sediment Acoustics Experiment (SAX99), were analyzed to determine the penetration processes. The two main candidate processes are refraction and scattering. Wave front coherence and direction of arrival, as well as sound speed measurements, provide important clues concerning the processes involved. The sound was projected by a wideband sound source carried on a re-

motely operated vehicle (ROV) in the band 10 to 100 kHz. Coherent and incoherent processing methods were explored. [Work supported by the Office of Naval Research, Ocean Acoustics.]

9:15

2aUW6. Subcritical acoustic scattering across a rough fluid–solid interface—laboratory experiment and perturbation theory model. Garfield R. Mellema (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada), Terry E. Ewart, and Kevin L. Williams (Univ. of Washington, Seattle, WA 98105)

Several mechanisms have been proposed by which an incident acoustic wave could be coupled into the ocean floor at angles below the compressional critical angle. Recent observations of acoustic transmissions from water into sediment with arrival times and amplitudes inconsistent with the refractive compressional path have been interpreted as the excitation of a Biot slow wave in the sediment. Another hypothesis attributes the observed signals to scattering at the rough water-sediment interface. A third entails scattering of evanescent waves by volume inhomogeneities in the sediment. The existence of multiple hypotheses, each of which could account for the received energy, invited further investigation. A series of well-controlled laboratory measurements were made and compared with the results of a first-order perturbation theory model in order to evaluate the accuracy of the rough interface hypothesis. The experimental measurements of transmission through adjacent flat and rough interfaces show good agreement with the results of the numerical model, giving clear evidence of subcritical penetration in an environment incompatible with the requirements of the Biot slow wave and volume scattering hypotheses.

9:30–9:45 Break

9:45

2aUW7. High-frequency measurements of absorption and dispersion in a sandy sediment. Darrell R. Jackson, Kevin L. Williams, Dajun Tang, and Eric I. Thorsos (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

As part of the SAX99 experiment, a buried hydrophone array was deployed together with a movable tower with attached sources covering the frequency range 11–50 kHz. With the tower placed to provide incident grazing angles well above the critical angle, this system was used to obtain data from which sediment sound speed and absorption were determined. The sound-speed data exhibit significant dispersion, while the absorption data show an approximate linear frequency dependence. When these data are combined with data at other frequencies from the same site, the dispersion and absorption are found to be consistent with causality and with the Biot model [Work supported by ONR.]

10:00

2aUW8. Velocity dispersion in water-saturated granular sediment. Robert D. Stoll (Lamont-Doherty Earth Observatory of Columbia Univ., Palisades, NY 10964, stoll@ldeo.columbia.edu)

When the Biot theory is applied to problems in sediment acoustics, significant velocity dispersion is predicted over a specific frequency range that depends on the pore-fluid mobility (permeability). This dispersion, and the associated nonlinear variation of attenuation with frequency, affects the reflection and transmission characteristics of the water–sediment interface, particularly at low grazing angles. Low-frequency velocity measurements taken during the ONR SAX99 experiments, when considered together with the high-frequency data, verify the existence of this dispersion. This result has important implications in the choice of critical angle and the interpretation and extrapolation of velocity and attenuation data taken at high frequencies. [Work supported by ONR, Code 321OA.]

10:15

2aUW9. Acoustic wave propagation in granular media. Masao Kimura (Dept. of Geo-Environ. Technol., Tokai Univ., 3-20-1 Orido, Shimizu, Shizuoka, 424-8610 Japan, mkimura@scc.u-tokai.ac.jp) and Nicholas P. Chotiros (Univ. of Texas, Austin, TX 78712)

The frame bulk and shear moduli are important in analyzing the acoustic wave propagation in marine sediments. The frame bulk and shear moduli of marine sediment are related to the longitudinal and shear wave velocities in dry marine sediments. The measured results of the longitudinal wave velocities in dry glass beads with four different grain sizes, but with almost the same porosity have been reported. The measured results showed that the longitudinal wave velocity largely increases, as the grain size increases. This result could not be explained using the classical contact theory, such as the Hertz–Mindlin or Digby model. In this study, we use the longitudinal wave velocities of dry glass beads in air and in vacuum to investigate the effects of the air, as the pore fluid. The measured results show that the longitudinal wave velocities in vacuum are smaller than that in air, and the velocities in vacuum become almost constant, even though the grain sizes are different.

10:30

2aUW10. Sediment acoustic backscattering during SAX99: Measurements and models. Kevin L. Williams, Darrell R. Jackson, Dajun Tang, and Eric I. Thorsos (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

During the SAX99 experiment, acoustic backscattering measurements were made at frequencies from 20 to 300 kHz as a function of grazing angle. The results from these acoustic measurements will be presented and compared with backscattering models that use the environmental measurements of other SAX99 researchers as input. In the 20–50-kHz range these comparisons indicate that surface roughness plays a dominant role in acoustic backscattering with a very distinctive reduction in backscattering at grazing angles above the critical angle of the sediment. Above 50 kHz this critical angle feature is less evident. Possible reasons for this change with frequency will be discussed. The backscattering models used here were originally developed for frequencies from 10 to 100 kHz. SAX99 data give some indication that further modeling is needed above 100 kHz. [Work supported by ONR.]

10:45

2aUW11. Biological and hydrodynamic control of seafloor roughness: Implications to high-frequency acoustic scattering. Michael D. Richardson, Kevin B. Briggs (Marine Geosciences Div., Naval Res. Lab., Stennis Space Center, MS 39529-5004), and Kevin L. Williams (Univ. of Washington, Seattle, WA 98105-6698)

Biological and hydrodynamic processes can both create and destroy seafloor microtopography. As part of the SAX99 experiments, natural and artificial temporal changes in seafloor roughness were monitored acoustically and quantified using bottom stereo photographs. Feeding activities of benthic megafauna and fish destroyed large-scale roughness features generated by ocean surface gravity waves within a period of weeks to months; whereas, fine-scale roughness created by raking the seafloor decayed to background levels within 24 h. The effects of fine-scale roughness increased acoustic scattering centered at one-half the acoustic wavelength (a Bragg wavelength of 2 cm) by 12–18 dB in artificial manipulations of the bottom. These changes were restricted to roughness that was oriented predominantly orthogonal to the incident acoustic waves. Alternatively,

seafloor roughness generated by ocean surface gravity waves had wavelengths of 50–100 cm and wave heights of 10–15 cm. These predictable large-scale roughness features should, by analogy, dramatically increase scattering at lower acoustic frequencies (near 1–2 kHz) and decay within weeks to months after storm events. [Work supported by ONR.]

11:00

2aUW12. High-frequency, low-grazing-angle backscattering from a sandy bottom. Charles F. Greenlaw, D. Van Holliday, and Duncan E. McGehee (Bae Systems, 4669 Murphy Canyon Rd., San Diego, CA 92123)

Backscattering measurements from a sandy bottom off East Destin, FL were obtained using a multi-frequency echosounder attached to a pan/tilt head on a bottom-mounted tower during the Sediment Acoustics Experiment, SAX99. Usable data were obtained over the frequency range from 265–1850 kHz, spanning the transition from roughness-controlled to grain-size-controlled scattering, at grazing angles around 15. Field results will be discussed and compared to laboratory measurements. [Work supported by ONR Code 322 BC.]

11:15

2aUW13. The small slope approximation for acoustic scattering from a rough interface between a fluid and a fluid-saturated porous solid. Cristina Fetche, Shira L. Broschat (School of Elec. Eng. and Computer Sci., Washington State Univ., P.O. Box 642752, Pullman, WA 99164-2752), and Taiqian Yang (Central Washington Univ., Puyallup, WA 98374)

The small slope approximation, introduced by A. G. Voronovich in the 1980s, gives a systematic expansion in terms of a generalized surface slope. Previous numerical studies indicate that the lowest-order term provides a reasonably accurate approximation for a number of different problems of practical importance. In this work we consider the small slope approximation for acoustic scattering from a rough interface between a fluid and a fluid-saturated porous solid in the context of Biot theory. Biot theory is of continued interest in the study of sediment acoustics; it predicts a slow compressional wave in the fluid-filled porous medium in addition to the conventional compressional and shear waves found in an elastic solid. The lowest-order small slope approximation is derived for Biot theory, and an expression for the bistatic scattering cross section is presented for plane-wave scattering from a randomly rough interface between a fluid and a fluid-saturated porous solid. Numerical results for the scattering strength are discussed for both the small slope and perturbation methods. [Work supported by ONR.]

TUESDAY AFTERNOON, 5 DECEMBER 2000 NEWPORT ROOM SOUTH, 1:30 TO 3:00 P.M.

Session 2pAA

Architectural Acoustics: Building Renovation and Sound Scattering and Absorption

Courtney B. Burroughs, Chair
Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

Contributed Papers

1:30

2pAA1. Acoustics of a remodeled historic building. Ballard W. George (Environtech Consultants, 1367 Bobolink Circle, Sunnyvale, CA 94087)

An acoustical study was performed for a historic San Francisco industrial building being remodeled for office, restaurant, and exhibit use. Challenges involving restoration ground rules included open beam construction with no use of ceilings. Flooring is proposed to be hardwood. Damped plywood was recommended to be used in flooring assemblies to enhance airborne and impact sound insulation performance. The building is across the street from the newly opened Pac tel baseball park for the San Francisco Giants. Accordingly, exterior noise sources included both automobile and bus traffic and the various sounds associated with baseball games. Sound measurements were taken at the 3 Com Park on Candlestick Point as an aid in predicting the sound levels associated with the new park. The acoustical study addressed the restaurants and exhibit spaces on the lower floors, the separating flooring assemblies between each level, and the sound ''delivered'' by ducting, and otherwise transmitted, to the future tenant spaces. Comments are given on the ''soundscape'' considerations associated with the various spaces of the lower floors of the building.

1:45

2pAA2. Room boundary absorption coefficient measurements in a small room. Jerald R. Hyde (JRH Acoustics, Box 55, St. Helena, CA 94574)

During the renovation of an 1890s Victorian building, the author's high ceiling office became entirely empty, the rectilinear boundaries consisting of wood, plaster, and glass surfaces. With a volume of 86 m^3 (3030 ft^3) the empty room had reverberation times (RT) ranging from 1 sec at low frequencies to over 2-1/2 sec at high frequencies. Boundary absorption coefficients were derived from RT data for the empty room's surfaces with the greatest absorption occurring in the 125-Hz octave band. Two different (weight) carpets of approximately 8.5 m^2 (92 ft^2) were then added to the floor and their absorption values calculated. Absorption coefficient results will be presented along with a study of the derived coefficient values of each of the component surfaces. Similar ''boundary only'' results from an ''empty'' 500-seat theater will also be shown.

2:00

2pAA3. Scale-model MLS measurements of edge diffraction. Rendell R. Torres (Chalmers Rm. Acoust. Group, Chalmers Univ. of Tech., SE-412 96 Gothenburg, Sweden) and Michael Vorländer (Tech. Univ. of Aachen, Aachen, Germany)

To better understand edge diffraction in room acoustics, scattering is measured from 1:10 scale models of reflector arrays and a stage house. The source is an 11-cm dodecahedron with tweeter elements having an effective frequency range of 600–6000 Hz. Maximum-length sequences (MLS) are used, which have various advantages over spark sources for scale models. The main three reflector arrays have approximately the same equivalent area: an array of five long rectangular panels, a 7-by-5 array of smaller square panels, and a single solid panel. Measurements are taken along a line under the reflector arrays and repeated for different array sizes and coverage. The stage-house measurements include a stepped reflector array and receiver positions shadowed from the receiver. Results compare well with previous work [R. W. Leonard *et al.*, J. Acoust. Soc. Am. **36**,

2328–2333 (1964)], and some practical rules of thumb are discussed. [Work performed as guest of the Inst. fuer Techn. Akustik, RWTH-Aachen, Germany; support provided by the Axson Johnson Foundation, Sweden.]

2:15

2pAA4. Pressure-amplitude reflection coefficients of active impedance tube terminations and their relationships to energy-related quantities of adjacent sound fields. Timothy W. Leishman (Dept. of Phys. and Astron., Brigham Young Univ., Provo, UT 84602, tim_leishman@byu.edu)

Active boundary surfaces intended to control reverberation or other characteristics of enclosed sound fields have often been investigated using impedance tubes. Unfortunately, the significance of pressure-amplitude reflection coefficients produced by these surfaces and their relationships to acoustic energy quantities in enclosed spaces has remained obscure. This paper accordingly presents a straightforward analysis of an actively terminated impedance tube to provide needed clarification. It demonstrates that increases in reflection coefficient moduli at the termination do not necessarily produce proportional increases in total enclosed energy in the tube or energy flux toward the enclosed field. Very large reflection coefficients (i.e., with moduli much greater than unity) produce both total enclosed acoustic energies and vector magnitudes of net active intensity approaching those present under an anechoic boundary condition. Large relative changes in total enclosed energy occur only when active termination reflection coefficient moduli roughly approach unity. Energy-related effects of actively produced reflection coefficients are generally frequency-dependent functions of distinct enclosed spaces and acoustic fields generated by primary sources in these spaces. Hence, correlations between reflection coefficient moduli and energy quantities of adjacent sound fields are not necessarily simple or consistent. They depend upon important physical properties of acoustic spaces, sources, and source arrangements.

2:30

2pAA5. Normal-incidence sound absorption measurement of a perforated panel using a sound-intensity technique: Theoretical versus experimental. Mohamad N. Dimon, Tharek A. Rahman, Ahmad K. Said (Elec. Eng. Faculty, UTM, 81310, Skudai, Johor, Malaysia), M. N. Ibrahim (UTM, 81310, Skudai, Johor, Malaysia), and M. Y. Jaafar (MARDI, Serdang, Selangor, Malaysia)

A direct-piercing carved-wood panel (DPCWP) is a form of wood carving that allows air to pass through the panel easily. In Malaysia, DPCWPs are available with various perforation ratios, patterns, and elements, and are typically used as part of a wall partition for an enclosed room. It helps in providing natural ventilation, besides it is believed to be able to act as a sound absorber. Circular perforated panels are the simplest form of DPCWP available. It was found out that normal-incidence sound-absorption performance of perforated panels can be measured reliably using a sound-intensity measurement technique. In this paper, normal-incidence sound-absorption measurement of a perforated panel using fixed-point scanning will be discussed. Close correlations were obtained from 250 Hz to 4 kHz between theoretical and measured sound absorption. Factors contributing toward sound intensity measurement reliability, primarily through its sound quality indicators such as $F1$, $F2$, $F3$, and $F4$, and the effect of measurement distance will be thoroughly discussed. This finding suggests that the sound intensity measurement technique can be used to measure normal-incidence sound-absorption performance of DPCWP reliably.

2:45

2pAA6. Acoustic properties of a triple-leaf membrane. Kimihiro Sakagami, Masayuki Morimoto (Environ. Acoust. Lab., Kobe Univ., Nada, Kobe, 657-8501 Japan, saka@kobe-u.ac.jp), and Masakazu Kiyama (Nippon Sheet Glass Environment Amenity Co., Ltd., Osaka, 541-0045 Japan)

Acoustic properties of a triple-leaf membrane, consisting of two impermeable and one permeable leaves, are theoretically analyzed under a simple condition in which a plane wave is incident normally. The triple-leaf membrane shows different acoustic properties according to where the permeable leaf is placed. When the permeable leaf is placed in the cavity between two impermeable leaves, its absorptivity shows two mass–air–mass resonance peaks, one by the first and second leaves, and the other by the first and third leaves, provided that the permeable leaf is of sufficient flow resistance and mass. These peaks can be merged to produce one larger peak so that the structure can offer higher absorptivity over a wider frequency range, when the masses of leaves and the cavity depths are adequately adjusted. A parametric study shows that the sound absorption and insulation performances of a triple-leaf membrane of this type are better than those of a double leaf with an air cavity only, though a double leaf with an absorbent in its cavity is still the most efficient. A method to optimize the system parameters to obtain maximum absorptivity is also suggested. Another arrangement, with the permeable leaf as the first leaf, is also explored.

TUESDAY AFTERNOON, 5 DECEMBER 2000 PACIFIC SALON C, 1:00 TO 2:05 P.M.

Session 2pABa

Animal Bioacoustics: Distinguished Lecture—Ultrasonic Eyeglasses for the Blind

Whitlow W. L. Au, Chair
Hawaii Institute of Marine Biology, P.O. Box 1106, Kailua, Hawaii 96734

Chair's Introduction—1:00

1:05

2pABa1. Ultrasonic eyeglasses for the blind. Leslie Kay (Spatial Sensing Lab., Sonicvision Ltd., P.O. Box 124, Russell, New Zealand)

There is an obvious need for blind persons to perceive their environment in the greatest detail possible. The bat has demonstrated the effective use of ultrasonic waves for its own purpose of perceiving its environment. Man now has his own air-sonar technology, in the form of eyeglasses mounted on the head, so as to perceive his environment. It may well be the best that can be achieved for helping blind persons by using their auditory channels to the cortex for spatial imaging and cognition of object space. This address describes the octave band CTFM air sonar with auditory coupling to the brain. The method of testing its physical performance is related to the human auditory perception of object space experienced by users. FFT analysis shows a distal resolution of approximately two wavelengths. However, behavioral analysis of a user performance suggests a greater neural ''resolution'' of object space is available. This resolution is demonstrated by the natural real-time perception of object-shape characteristics. The ''infinity'' of sonar object ''signatures'' are seemingly absorbed by the listener's neural system and eventually learned. The small nuances in the real-time change in sound structure as an object pose is altered are invariant. This requires an enormous memory.

TUESDAY AFTERNOON, 5 DECEMBER 2000 PACIFIC SALON C, 2:20 TO 5:10 P.M.

Session 2pABb

Animal Bioacoustics and Noise: Standards and Animals: Where Are We?

Ann E. Bowles, Chair
Hubbs-Sea World Research Institute, 2595 Ingraham Street, San Diego, California 92109

Invited Papers

2:20

2pABb1. The development of human standards to limit noise-induced hearing loss: A good model for animals? Ann E. Bowles (Hubbs-Sea World Res. Inst., 2595 Ingraham St., San Diego, CA) and Daniel L. Johnson (Interactive Acoustics, Provo, UT)

Over the past 50 years, the ASA and other organizations have promulgated standards to protect human hearing. We are evaluating these standards as a model for similar standards for animals. First, we must eliminate noise at frequencies that could have little effect on hearing. In the case of humans, A-weighting has been selected for most applications. The choice was determined partly by convenience (readily available circuitry, computationally simple calculations for weighted values) and partly by its robustness in fitting empirical data. Second, allowable exposures must take into account exposure time. For humans, the time-intensity tradeoff is calculated on an equal-energy basis that presumes underlying linear functions (the basis of the 3-, 4-, and 5-dB exchange values used by various organizations), again to simplify computation. For high-amplitude, brief transients (impulse noise), which produce nonlinear auditory responses, an absolute upper limit of exposure is used. This scheme has proved adequate to protect crucial human hearing capabilities (mainly the understanding of speech), but may not be conservative enough for animals, which use their hearing in self-defense, for navigation, etc. It is interesting to note that schemes for animals can be developed without the computational simplicity that historically drove many decisions in the human case.

2:40

2pABb2. A preliminary study of loudness at frequencies of 5 to 120 kHz based on whistle response time (RT) in a dolphin. Sam Ridgway and Donald Carder (Marine Mammal Prog., Space and Naval Warfare Systems Ctr., 49620 Beluga Rd., San Diego, CA 92152-6266)

Stebbins (1966) showed that auditory RT of monkeys could be used to measure loudness. Such studies have been replicated in other animals and humans. Although Ridgway *et al.* (1991) reported that dolphin auditory RT varied with stimulus intensity, curves showing variation in RT with frequency and amplitude were not presented. Choice RT was measured in a female bottlenose dolphin responding to 300 msec bursts of clicks (30 clicks at 10 ms ISI) with a burst of pulses, and to 300 msec tones with a whistle. Responses were received with a hydrophone (B&K 8103), digitized, and stored on disk. The dolphin responded correctly (whistled to the tone, pulsed to the pulses) on more than 98% of presentations. Here we only consider tone presentations. Frequencies, presented randomly, were 5, 10, 20, 40, 60, 80, 100, and 120 kHz and ranged from 85 to 135 dB *re* 1 Pa rms. At 10 kHz, median RT ranged from 240 msec at 135 dB to 410 msec at 85 dB. At 100 kHz, median RT ranged from 260 msec at 135 dB to 330 msec at 85 dB. Studies must be done over a larger range of amplitude and frequency and with precise control of background noise.

3:00

2pABb3. Statistical requirements to demonstrate subtle impacts of noise on passerine bird populations. Don Hunsaker, II (Hubbs Sea World Res. Inst., 2595 Ingraham St., San Diego, CA 92109)

We are conducting a study on the effects of noise on California gnatcather (CAGN) and least Bell's vireo (LBV) populations at military installations in San Diego, CA. In order to improve the statistical power to be able to accurately measure subtle effects of noise on the birds, we have had to expand the size of the study considerably. We have mapped noise levels and reproductive success in the study area. This provides an estimation of the noise covariate at all the sampled nests and enables us to characterize the magnitude of effects of noise on reproductive success by estimating the strength of correlation between noise and nesting success endpoints. Our analysis determined that 60 nesting pairs would provide data to detect a 20% difference at the 5% significance level with 80% power. This ability to measure only gross impacts was unacceptable to us. The design now includes 5 years to sample 300 nests of CAGN and 800–1000 nests of LBV for us to achieve the power to assess these subtle effects at the 5% level.

3:20

2pABb4. Assessing ocean noise: A conservationist approach. Joel R. Reynolds and Michael D. Jasny (Natural Resources Defense Council, 6310 San Vicente Blvd., Ste. 250, Los Angeles, CA 90048)

The regulation of ocean noise is an evolving process that involves representatives of the environmental community as well as scientists and regulators. This paper describes a three-stage procedure, based on the precautionary principle, that is currently being used by environmentalists to identify acoustic projects with potentially significant biological impacts: (1) a threshold stage of review, in which factors such as the ecological significance of the operation area, the intensity, frequency, and duration of the source, the mitigation offered by project applicants, and the cumulative and precedential impacts of the project are considered; (2) a period of consultation with experts in the field; and (3) an assessment of the project according to the substantive, procedural, and evidentiary standards set forth by applicable statutes and regulations, including the National Environmental Policy Act, Marine Mammal Protection Act, and Endangered Species Act. In the third stage, projects are evaluated according to the integrity of their alternatives analysis, especially with regard to location and season; the consistency of their environmental assessment with the best available science; the significance of their reasonably anticipated cumulative impacts; their objectivity in identifying data gaps; and their commitment through mitigation to have the least practicable impact on the marine environment.

Contributed Papers

3:40

2pABb5. Masked temporary threshold shift (MTTS) in odontocetes after exposure to single underwater impulses from a seismic watergun. James J. Finneran (Marine Mammal Prog., Space and Naval Warfare Systems Ctr., 49620 Beluga Rd., San Diego, CA 92152-6266), Carolyn E. Schlundt, Randall Dear (Sci. Applications Intl. Corp., San Diego, CA 92110), Donald A. Carder, and Sam H. Ridgway (Space and Naval Warfare Systems Ctr., San Diego, CA 92152-6266)

A behavioral response paradigm was used to measure masked underwater hearing thresholds in a bottlenose dolphin (*Tursiops truncatus*) and white whale (*Delphinapterus leucas*) before and after exposure to single underwater impulses produced by a seismic watergun. Hearing thresholds were measured at 0.4, 4, and 30 kHz using a vocal response technique and a modified version of the Method of Free Response. A pre-exposure threshold was generally measured 34 min before presentation of the impulse. Post-exposure threshold measurements began within 2 min of exposure to the impulse and continued for approximately 20 min following exposure, in order to track recovery. A masked temporary threshold shift (MTTS) was defined as a 6-dB or larger increase in threshold compared to the corresponding pre-exposure measurement. The MTTS test paradigm will be discussed and preliminary MTTS and behavioral response data presented and compared to existing TTS data for odontocetes. [Work supported by ONR.]

3:55

2pABb6. Underwater temporary threshold shifts in pinnipeds: Tradeoffs between noise intensity and duration. Ronald Schusterman, David Kastak, Brandon Southall, and Colleen Kastak (Long Marine Lab., Univ. of California, Santa Cruz, 100 Shaffer Rd., Santa Cruz, CA 95060, rjschust@cats.ucsc.edu)

Temporary threshold shifts (TTSs) were obtained in one female California sea lion (*Zalophus californianus*), one male harbor seal (*Phoca vitulina*), and one female northern elephant seal (*Mirounga angustirostris*). Prior to noise exposure, base-line hearing thresholds were obtained at 2500 or 3530 Hz using behavioral techniques. During noise exposure, the subjects were stationed in an underwater noise field for either 20–25 or 50 total minutes. The noise field consisted of octave-band noise centered at 2500 Hz with exposure levels of either 65, 80, or 95 dB above the mean base-line thresholds [i.e., sensation level (SL)]. Postexposure thresholds were obtained immediately following noise exposure. The same testing and diving patterns were used in the absence of noise to control for the effects of diving. Hearing losses ranged from 0 to 28 dB, and there was substantial inter- and intraspecific variability in the amount of hearing loss following exposure to similar noise regimes. Recovery was also variable, ranging from minutes to hours, with complete recovery always occurring within 24 hours. We found threshold shifts of 5 dB at 65 dB SL and only

marginal increases in TTS with increasing noise intensity. However, increasing noise duration with constant noise exposure levels yielded greater shifts.

4:10

2pABb7. A comparison of behavioral and auditory brainstem response methods for examining hearing sensitivities in the harbor seal (*Phoca vitulina*). Lawrence F. Wolski, Rindy C. Anderson, and Ann E. Bowles (Hubbs-SeaWorld Res. Inst., 2595 Ingraham St., San Diego, CA 92109)

Auditory brainstem response (ABR) and behavioral methods (method of constant limits and staircase method) were used to generate in-air audiograms for an adult female harbor seal to examine relative sensitivities obtained using the two methods. Behavioral thresholds ranged from 0.250–30 kHz and indicated best sensitivity from 6–12 kHz (~12 dB *re*: 20 μPa). The staircase method was more sensitive than the method of constant limits, resulting in thresholds averaging 10 dB lower. ABR thresholds were recorded at 2, 4, 8, 16, and 22 kHz and showed a similar range of best sensitivity (8–16 kHz). ABR and behavioral thresholds differed by an average of 5 dB; ABR thresholds averaged 3.8 dB higher than behavioral thresholds from 2–8 kHz, but were 12 and 3 dB lower at 16 and 22 kHz, respectively. The lower ABR thresholds at higher frequencies may reflect the behavior of the seal during ABR testing as well as broadband characteristics of ABR test stimuli. These results agree with comparisons of ABR and behavioral methods performed in other recent studies and indicate that auditory evoked potential methods represent an accurate alternative to behavioral methods for estimating hearing range and sensitivity in wild animals. [Work supported by NASA: NAS1-20101.]

4:25–5:10

Panel Discussion

TUESDAY AFTERNOON, 5 DECEMBER 2000 PACIFIC SALON D, 1:00 TO 5:50 P.M.

Session 2pBB

Biomedical Ultrasound/Bioresponse to Vibration, Physical Acoustics and Acoustical Oceanography: Detection and Characterization of Bubbles, Acoustic Cavitation, and Associated Physical Effects II

Ronald A. Roy, Cochair
Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Timothy G. Leighton, Cochair
Institute of Sound and Vibration Research, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom

Chair's Introduction—1:00

Invited Papers

1:05

2pBB1. Characterization of measures of reference acoustic cavitation (COMORAC). Timothy G. Leighton (Inst. of Sound and Vib. Res., Univ. of Southampton, Highfield, Southampton SO17 1BJ, UK)

This paper describes a coordinated experiment involving a number of UK centers, with the aim of comparing a range of systems which measure cavitational activity. Sensors have been designed at a number of universities to produce quantitative outputs based on sonochemical (Bath and Southampton), calorimetric (Coventry), electrochemical and luminescent (Southampton) measures. In addition, acoustic sensors have been produced (Guys and St. Thomas Hospital Trust, and the National Physical Laboratory). These are all deployed to monitor controlled cavitation events, ranging from single bubble phenomena to a reference cavitation bath (produced by NPL). Design of these cavitation systems has involved not only the characterization of the sound field and the vessel, but also the specification of some reference cavitation liquid (NPL and Southampton). Comparison of the resulting measures of activity allows the feasibility of a scale for cavitation to be investigated.

1:30

2pBB2. Optical Mie scattering to observe single cavitation bubble dynamics. R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, rgholt@bu.edu)

It is often advantageous to know precisely the instantaneous response of bubbles to an external acoustic field. One technique to enable such observations consists of detecting the time-resolved optical scattering of an incident laser beam from a bubble. For visible laser light and bubble radii on the order of 1–100 microns, reflection, refraction, internal interference, and diffraction must all be taken into account, and the scattering so described is named after Gustav Mie. Physical principles and the mathematical development will be discussed. Several specific applications will be presented and the results discussed. Finally, a brief list of ''do's and don'ts'' for the potential user will be discussed.

1:55

2pBB3. When is a bubble a bubble?: The subtleties of cavitation threshold detection at MHz frequencies. Ronald A. Roy (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215)

Detecting the onset of cavitation induced by short-pulse ultrasound at MHz frequencies is tricky business, for one must be positioned to observe the genesis and evolution of a single micron-sized bubble lasting as short as one acoustic cycle. The observation is rarely accomplished directly. Rather, investigators often rely on indirect evidence of cavitation activity manifested by physical effects such as noise (a.k.a. passive cavitation detection), light emission, and mechanical damage. These techniques are noise limited, and investigators frequently measure a threshold for ''detectable'' cavitation. Another approach is to actively probe the medium with a high-frequency acoustic beam, and watch for anomalous backscatter associated with bubble activity. This technique, often termed active cavitation detection, can be quite sensitive, which is part of the problem; transient disturbances in the medium can be mistaken for bona fide cavitation activity. Such errors are often associated with, say, the motion of solid particles in a test liquid that is put into motion by acoustic radiation forces or entrained by acoustic streaming. In all cases, the ability to detect and monitor minute, spurious cavitation activity is limited by noise, clutter, or reverberation. These issues are discussed, examples given, and potential solutions proposed. [Work supported by DARPA.]

2:20

2pBB4. Acoustic spectrum and sonoluminescence as indicators of cell membrane disruption by ultrasound. Mark R. Prausnitz, Hector R. Guzman, and Stephen Cochran (Schools of Chemical and Biomed. Eng., Georgia Tech, Atlanta, GA 30332-0100)

Ultrasound has been shown to reversibly disrupt cell membranes and thereby drive drugs and genes into cells by a mechanism believed to involve cavitation. We have measured the number of molecules delivered into cells and levels of cell viability following exposure to a broad range of ultrasound conditions. Many cellular responses are possible, including uptake of up to millions of molecules per cell and cell viability ranging from 0 to 100%. Because these effects are believed to occur due to acoustic cavitation, we measured features of the acoustic spectrum known to be associated with cavitation and measured sonoluminescence light output during ultrasound exposures. We found that these measures of cavitation correlate with both the number of molecules taken up by cells and levels of cell viability. This observation strengthens the hypothesis that ultrasound's bioeffects are mediated by cavitation and suggests a noninvasive means to assess these bioeffects which could be used as part of a real-time feedback loop. [Work supported by the NSF, NIH, and Whitaker Foundation.]

2:45

2pBB5. Control of inertial cavitation in shockwave lithotripsy. Andrew J. Coleman (Med. Phys. Dept., Guy's and St. Thomas' Hospital, London SE1 7EH, UK, andrew.coleman@gstt.sthames.nhs.uk) and Mark D. Cahill (Univ. of Bath, Bath BA2 7AY, UK)

The role of inertial cavitation in stone comminution and tissue damage during extracorporeal shock-wave lithotripsy (ESWL) became appreciated several years after the successful introduction of ESWL into routine clinical use in the early 1980s. It is only relatively recently, however, that attempts have been made to enhance or suppress the cavitation field in lithotripsy by adapting the acoustic field generated by the shock-wave source. The single-cycle acoustic pulses used in commercial ESWL generate a cavitation field in which the bubbles may remain in an expanded phase for many times the pulse duration. Some control of cavitation can be gained by timing an additional pulse from a second source to coincide, for example, with the bubbles expanded phase [M. R. Bailey, Tech. Rep. No. ARL-TR-97-1, Univ. of Texas, Austin (1997)]. This can generate an earlier and more violent bubble collapse. A single source generating multiple-cycle pulses can, similarly, provide a means of controlling cavitation activity in lithotripsy. This approach has been examined using an electromagnetic-type shock-wave source to generate high-amplitude, multicycle pulses in water. By tailoring the pressure waveform specifically to control cavitation, it may be possible to refine medical ultrasound applications, such as ESWL, where cavitation plays a role.

3:10

2pBB6. Characterization of ultrasound contrast agent oscillations. Katherine Ferrara, Karen Morgan, James Chomas, and John Allen (Div. of Biomed. Eng., Univ. of California, Davis, One Shields Ave., Davis, CA 95616, kwferrar@ucdavis.edu)

Ultrasound contrast agents consist of a gas bubble surrounded by a shell, and are used in studies of tissue perfusion. It is currently difficult to differentiate contrast agent echoes from nonlinear components of tissue echoes. A goal of this research is to improve the understanding of individual contrast agent oscillations and the resulting received echoes. This, in turn, can lead to more sophisticated bubble detection techniques by designing pulsing schemes to elicit a specific bubble response. In this study, experimental and theoretical tools are uniquely combined to aid in the understanding of a bubble's response during insonation. A model is used for comparison, which is based on the equation developed by Rayleigh for an unencapsulated bubble, with additional terms to include the effects of shell elasticity and viscosity. The resulting radius-time curves are validated by direct comparison with the experimental results from a high-speed optical system capable of 100 million frames/s. One resulting signal processing strategy involves the transmission of a train of wide-band pulses with alternating phase which yields echoes with significant spectral differences, quantified by calculating the mean frequency of each echo. A new bubble detection technique is proposed based on these differences in mean frequency.

3:35–3:50 Break

3:50

2pBB7. Nonlinear scatter from the ultrasound contrast agent Sonazoid: Theory and numerical simulations. Lars Hoff (Dept. of Telecommunications, Norwegian Univ. of Sci. and Technol., N-7491 Trondheim, Norway)

A model for the contrast agent Sonazoid is presented. The model is based on a modified Rayleigh–Plesset equation, with the shell added as an incompressible viscoelastic solid. Different models for the shell's nonlinear stress–strain relationship are discussed. Shell elasticity and viscosity were estimated from experiments, and the model was used to simulate bubble radius and scattered sound. Pulse-inversion imaging was simulated. The amplitude was increased from 20 to 500 kPa, giving scattered pulses that changed from inverted to time-shifted versions of each other. This effect was best seen in the cross correlation of the responses. The nonlinear response was strongest for frequencies below bubble resonance. Compared to unshelled bubbles, Sonazoid required higher drive amplitudes to respond nonlinearly. The Sonazoid bubbles responded similarly to unshelled bubbles of smaller diameter. A subharmonic response was found in a limited frequency band and only for diameters above 3 μm. An amplitude threshold was found around 500 kPa, depending on diameter and frequency. This threshold is higher than for unshelled bubbles. A chaotic oscillation pattern was found for some diameter–frequency combinations, in agreement with published results. For some situations where the cw results predicted a subharmonic mode, this mode was not obtained from finite-length pulses.

4:05

2pBB8. Interaction of shock waves with cavitation bubble: Application of Ho:YAG pulse laser irradiation for cerebral thrombolysis. S. H. R. Hosseini (Shock Wave Res. Ctr., Inst. of Fluid Sci., Tohoku Univ., 2-1-1 Katahira, Aoba, Sendai, 980-8577 Japan, hosseini@ceres.ifs.tohoku.ac.jp), Takayuki Hirano, Osamu Onodera, and Kazuyoshi Takayama (Tohoku Univ., Sendai, Japan)

This paper reports on the application of underwater shock waves and vapor cavities for cerebral thrombolysis. The energy source was a *Q*-switched Ho:YAG laser (Nippon Infrared Industries Co., Ltd.) with 91 and 140 mJ/pulse energy measured at the end of a 0.60-mm-diam glass optical fiber, pulse duration of 200 ns, and wavelength of 2100 nm. The whole sequences of the shock-wave propagation and growth of the cavitation bubbles from the end of the optical fiber were visualized by high-speed photography using an Imacon 468 camera (Hadland Photonics Ltd.). The shock-wave generation was associated with the laser breakdown in the water. The laser interaction produced a microplasma and heated the liquid in front of the fiber. The plasma drove spherical shock wave in water, followed by the formation of a high-temperature vapor cavity. Collapse of the cavity produced a secondary spherical shock wave in water. Successive generation and collapse of the bubbles were visualized. Pressure histories were measured by PVDF needle hydrophones. The underwater shock waves and the vapor cavities were successfully applied for disruption of artificial thrombosis. The result clearly shows potential extension of this method to precise medical treatment for revascularization in neurosurgery.

4:20

2pBB9. Radiation and ultrasound-triggered bubble contrast agents based on superheated drops. Robert Apfel (Yale Univ., New Haven, CT 06520-8286)

Superheated drop compositions have been formulated for infusion in the body. Low doses of x rays or diagnostic ultrasound pulses are sufficient to nucleate the vaporization of these drops, producing bubbles that can serve as ultrasound contrast agents for imaging. Drugs can be carried in the drop material and can be released at a specific site where the radiation or ultrasound is focused. The current work describes the status of: (1) the optimization of the components that make up the composition, which include the superheated material, another liquid to adjust the degree of superheated, an encapsulant, and possibly a drug; (2) tests of the sensitivity of these drops to being triggered either by x rays or sufficiently intense pulses of ultrasound; (3) the analytic tools for characterizing the compositions and their triggering; and (4) the practical challenges to making this system suitable for clinical application.

4:35

2pBB10. Effect of dual-reflector lithotripter on stone fragmentation and cell damage. Dahlia L. Sokolov, Michael R. Bailey, and Lawrence A. Crum (Appl. Phys. Lab., 1013 NE 40th St., Seattle, WA 98105)

An electrohydraulic lithotripter uses an ellipsoidal reflector to focus shock waves on to a kidney stone. The shock wave generates a cylindrical cavitation field, $\sim$1 cm wide$\times$10 cm long, that has been implicated in both stone fragmentation and damage to healthy tissue during lithotripsy treatment. A dual-reflector lithotripter, consisting of two identical spark-gap lithotripters facing each other and firing simultaneously, creates a more localized cavitation field, $\sim$3 cm wide$\times$5 cm long (Sokolov, Berlin 1999). Such a field may increase the rate of stone fragmentation while mitigating damage to surrounding tissue. Breakage of model stones at the focus and hemolysis of red blood cells 3 cm from the focus were assessed for both conventional lithotripsy (CL) and dual-reflector lithotripsy (DRL). To equalize total energy input, the number of shots was halved from CL to DRL. Stones subjected to DRL were broken into several major fragments while stones subjected to CL remained intact except for some dust. Human blood was diluted to 3% hematocrit in degassed PBS and placed in acoustically transparent sample tubes. There was no statistical difference in percent hemolysis between CL ($5.12\pm1.01\%$) and DRL ($5.39\pm0.57\%$). [Work supported by NIH DK43881 and NSF.]

4:50

2pBB11. Cavitation enhanced acoustic hyperthermia. Patrick Edson, R. Glynn Holt, and Ronald A. Roy (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215)

The dominant mechanisms through which cavitation activity may lead to enhanced heating in tissue and tissue-like media with high-intensity focused ultrasound remain unclear. Previously, the authors have reported the results of experimental [J. Acoust. Soc. Am. **104** (A) (1998)] and numerical [*ibid.* **106** (A) (1999)] investigations which sought to understand the relationship between parameters relevant to the cavitation field, dominant heating mechanisms, and the magnitude of the enhanced heating effect. The model has since been extended to better account for heating from acoustic re-radiation. Results of expanded numerical simulations which incorporate bubble related heating source terms in an existing FDTD code designed to calculate temperature response as a function of the incident acoustic field are reported. Experiments aimed at manipulating the cavitation field to facilitate reproducibility and control are also reported. [Work supported by DARPA.]

5:05

2pBB12. Hepatic vasculature flow monitoring through the ultrasonic interruption of contrast agents. R. T. Rhee, J. B. Fowlkes, J. M. Rubin, and P. L. Carson (Dept. of Radiol., Univ. of Michigan Medical Ctr., Ann Arbor, MI 48109)

A technique has been developed to assess differential liver perfusion due to hepatic artery versus portal vein. This technique uses a continuous intravenous infusion of contrast agent and transcutaneous ultrasound to control the flow of contrast agent in selected vessels. A 1-in diameter transducer controlled the flow of perfluorocarbon-filled contrast agent in the abdominal aorta. It has been previously shown that acoustic amplitudes required to manipulate contrast agent are well within FDA limits. Pulsed ultrasound 10-cycle tonebursts at 2.25 MHz and 2.7% duty cycle were applied to the abdominal aorta for 30 sec to clear contrast agent from the hepatic artery, portal vein and liver parenchyma. The interruption transducer was then turned off, creating a sharp leading edge of contrast

agent, with risetimes of approximately 0.1 sec measured in the hepatic artery. Harmonic grayscale and spectral Doppler images were obtained on a Toshiba Powervision 8000. The sharp contrast agent leading edge allows differential measurements of hepatic artery and portal vein arrival times. This time differential was measured at approximately 10 sec. These results indicate that image-guided insonification of the abdominal aorta may substantially improve discrimination of the hepatic arterial and portal venous supply in hepatic imaging.

5:20

2pBB13. Volume flow measurement using Doppler RF decorrelation. Theresa A. Tuthill, J. Brian Fowlkes, Jonathan M. Rubin (Dept. of Radiol., Univ. of Michigan Medical Ctr., Ann Arbor, MI 48109, tuthill@umich.edu), and Anne L. Hall (General Electric Medical Systems, Milwaukee, WI)

A technique for monitoring volumetric flow was developed by combining standard Doppler measurements with amplitude decorrelation. The proposed algorithm uses current clinical ultrasound technology and is independent of vessel orientation. However, to accurately measure the ultrasound signals' temporal decorrelation in blood, a contrast agent was added to increase the scattering signal strength. The algorithm uses standard frequency shifted Doppler signals, steered at two different angles, to determine the in-plane velocities. For the same firing lines, the corresponding temporal amplitude decorrelation is related to total velocity. The velocity normal to the imaging plane can then be extracted and integrated to compute total volume flow. The algorithm was tested in a flow tank using a GE Logiq 700 clinical scanner with a 7.5 MHz linear array. The output power was set at the lowest level to reduce effects from additional decorrelation due to acoustic radiation force. RF data were collected as in-phase and quadrature components. The tube flow results showed a parabolic velocity profile that scaled with volume flow. For transducer angles close to perpendicular to the vessel, the average calculated volume flow error was less than 25%. These preliminary results demonstrate the efficacy of our technique in measuring blood volume flow.

5:35

2pBB14. The ''underwater'' sounds from the impact of drops in superfluid ^{4}He. Antonios Apostolou (Grad. Prog. in Acoust., The Pennsylvania State Univ., State College, PA 16802), Jin H. So, Zhiqu Lu, J. Gladden, Eliza S. Bradley, and J. D. Maynard (The Pennsylvania State Univ., University Park, PA 16802)

When a drop of a liquid impacts onto a liquid surface a crater is formed, and the cavity dynamics that follow determine whether or not a bubble will be entrained. The entrainment of an air bubble following the impact of a raindrop onto the sea provides the predominant mechanism for the production of the underwater noise from rain. Experiments in water have indicated one particular entrainment region in a dimensionless parameter-space. This talk describes an ongoing research effort to study the phenomenon on a more fundamental basis using, instead of water, the pure and highly characterized liquid helium. The impact of liquid helium drops on a liquid helium surface is monitored by employing acoustic and optical techniques. Bubble entrainment in superfluid ^{4}He, as well as the transition to nonentrainment, are observed, and the collected data indicate a new entrainment region. In addition, high-speed images of the crater motion reveal novel crater dynamics for the nonentrainment case. [Research supported by National Science Foundation DMR 9801844 and the Office of Naval Research.]

TUESDAY AFTERNOON, 5 DECEMBER 2000 PACIFIC SALON A, 1:10 TO 3:15 P.M.

Session 2pEA

Engineering Acoustics: Acoustic Measurements and Materials Characterization

R. Lowell Smith, Chair
1749 Linbrook Drive, San Diego, California 92111-7112

Chair's Introduction—1:10

Contributed Papers

1:15

2pEA1. Evolutionary strategy algorithm for a complete characterization of porous materials using a standing wave tube. Youssef Atalla, Raymond Panneton (GAUS, Mech. Eng. Dept., Univ. of Sherbrooke, QC J1K 2R1, Canada, Raymond.Panneton@gme.usherb.ca), and Franck Sgard (LASH/DGCB URA CNRS 1652, ENTPE, France)

A widely used model for describing the attenuation of an acoustical wave propagating in a rigid open-cell porous material is the Johnson–Champoux–Allard (JCA) model. This model is based on five macroscopic parameters describing the porous medium: flow resistivity, porosity, tortuosity, viscous and thermal characteristic lengths. Simultaneously with the development of direct methods for measuring these five parameters, defining and solving inverse problems based on an artificial intelligence approach appears to overcome some of the limitations of the direct methods. In this work, an application of the evolutionary strategy (ES) algorithm for the estimation of the five parameters of a porous material from simple acoustical measurements is presented. First, a number of numerical tests are performed and the results are used like *a priori* knowledge of how to set up an evolutionary algorithm to solve such a difficult problem in the shortest time. In the second step, the final setup of the evolutionary algorithm is applied for evaluation of the five parameters from experimental measurements. The method seems practical and promising since it is based on simple acoustical measurements and avoids using complicated and unreliable measurement setups.

1:30

2pEA2. Comparison of techniques for measuring the acoustic properties of porous materials. C. Walter Beamer and Ralph T. Muehleisen (Dept. of Civil, Environ., and Architectural Eng., Univ. of Colorado, Boulder, CO 80309)

There is a large variety of methods for measuring the acoustic properties of porous materials. Of particular interest to many acoustics researchers are the characteristic impedance and propagating wave number of the porous materials. The standing wave impedance tube and two-microphone method of measuring those properties have both become standards. A number of extensions of the two-microphone method have been developed. Song and Bolton recently published a transfer matrix method utilizing four microphones [J. Acoust. Soc. Am. **107**, 1131–1152]. Comparisons of the tradeoffs and errors inherent in the major methods for the

measurement of the acoustical properties of porous materials are presented. [Work supported by the Office of Naval Research.]

1:45

2pEA3. Resonant acoustic measurement of the effective diffusion coefficient of gases through soil. Richard J. Schuhmann (Dept. of Civil & Environ. Eng., Penn State Univ., 212 Sackett, State College, PA 16802, rxs34@psu.edu), Matthew V. Golden, and Steven L. Garrett (Penn State Univ., State College, PA 16804)

Recent interest has developed within the environmental community in measuring the effective diffusion coefficient (D_e) of gases through soil. An electroacoustic system will be described that uses a plane-wave resonator and a phase-locked-loop automatic resonance frequency tracker. The gas analyzer is simple and can be built for about $100 from readily available plumbing parts and inexpensive transducers. A soil core is fitted into the top of a valved sample chamber and the resonator below is charged with a pure ''tracer'' gas (e.g., He, SF_6), then the valve opened. The rate at which the tracer gas is replaced by air within the resonator is controlled by the soil's D_e. The resonant frequency is maintained by a phase-locked-loop that compares the signal from an electret microphone with a square wave reference generated by the same monolithic function generator (XR 2206) that excited an electrodynamic loudspeaker with a sinusoidal current. The resonator temperature is monitored using a monolithic IC temperature sensor (AD 592) providing an output current of 1 μA/K. The mean molecular weight of the gas mixture in the resonator is directly determined in real-time from the ratio of the absolute temperature to the square of the fundamental acoustic resonance frequency. [Work supported by ONR.]

2:00

2pEA4. Relationship between the diffraction constants, the reciprocity parameter, and the mutual radiation impedance of a pair of compact radiators: An improved Pritchard approximation. Steven R. Baker (Phys. Dept., Code PH/Ba, Naval Postgrad. School, Monterey, CA 93943) and Clyde L. Scandrett (Naval Postgrad. School, Monterey, CA 93943)

By the principle of reciprocity, the mutual radiation impedance matrix of an acoustical array in a stationary fluid must be symmetric. However, use of the classical ''Pritchard approximation'' to compute array mutual (mechanical) radiation impedance ($\mathbf{Z}_{\text{mut}} = j\ \text{Re}[\mathbf{Z}_{\text{self}}]e^{-jkd}/kd$; $e^{j\omega t}$ assumed) [R. L. Pritchard, J. Acoust. Soc. Am. **32**, 730–737 (1960)] leads to a matrix which is not symmetric for nonidentical radiators. Also, it is restricted to low frequencies ($ka \ll 1$). By a simple application of the principle of reciprocity, it is shown that the mutual mechanical radiation impedance of a pair of reasonably compact radiators in an otherwise unbounded stationary fluid is expressible in terms of their (complex) diffraction constants $\mathbf{D}_1$ and $\mathbf{D}_2$ and the (complex) spherical wave reciprocity parameter $\mathbf{J}_s$ as $\mathbf{Z}_{\text{mut}} = S_1 S_2 \mathbf{D}_1 \mathbf{D}_2 / \mathbf{J}_s$, where S is the radiator surface area, $\mathbf{J}_s = 2d/j\rho f$, d is the separation between radiator acoustic centers, ρ is the fluid density, and f is the frequency. (For mutual *acoustical* radiation impedance, simply drop the factors S_1 and S_2.) For spherical radiators of radius a, $\mathbf{D}_n = e^{jka_n}/(1+jka_n)$. Note, this newly derived form for $\mathbf{Z}_{\text{mut}}$ is guaranteed to be reciprocal. Also, for small ka, it reduces to $S_1 S_2 / \mathbf{J}_s$, which is equivalent to the classical Pritchard approximation, corrected for nonidentical radiators.

2:15

2pEA5. Measurements of orifice impedance with grazing flow. Nolan S. Dickey, Ahmet Selamet (Dept. of Mech. Eng., The Ohio State Univ., Columbus, OH 43212, dickey.21@osu.edu), and Mehmet S. Ciray (Arvin Industries, Inc., Columbus, IN)

Perforated interfaces are commonly used in silencers for automotive applications. In these silencers, the acoustic behavior of the perforations, typically described in terms of a perforate impedance, can have a significant effect on silencer performance. The perforate impedance, in turn, is affected by the mean flow in the silencer. The focus of the present study is mean flow tangential to the perforate face (grazing flow), which is often predominant. In most cases, the perforate impedance cannot be determined analytically, and therefore experimental data and empirical models are necessary. Yet, experimental data suitable for use in models of perforated tube silencers are relatively scarce. The objective of the present investigation is to measure the impedance of perforated plates subjected to grazing flow. A description is given of the experimental apparatus, which has been designed for measurements with flow conditions representative of those in an automotive system. Data are presented for plates with square-edged circular orifices of varied dimensions. The results are reduced in terms of scaling parameters suitable for use in empirical perforate models.

2:30

2pEA6. Effects of nonuniform stress on the properties of lead zirconate titanate. Robert S. Janus, James M. Powers, Frank A. Tito, Mark B. Moffett, and Michael D. Jevnager (Naval Undersea Warfare Ctr., Newport, RI 02841, janusrs@npt.nuwc.navy.mil)

The variations in the piezoelectric constant, d_{33}, and the dielectric constant, $K_{33}{}^T$, for Navy type III lead zirconate titanate piezoelectric ceramic under nonuniform compressive mechanical stress have been experimentally evaluated. Peak compressive stress levels as high as 290 MPa (42 kpsi) with electric field intensities up to ± 790 kVrms/m (± 20 Vrms/mil) were applied to a stack of piezoelectric ceramic bars. The maximum d_{33} and stress-averaged $K_{33}{}^T$ values measured were 159% and 135% larger, respectively, than the corresponding catalog values specified at low field and without a mechanical bias. The electrical losses of the material measured under the above conditions are also reported. [Work supported by the Space and Naval Warfare Systems Command.]

2:45

2pEA7. Equivalent circuit representation of Poisson-coupled electroacoustic modes. R. Lowell Smith (Consultant, 1749 Linbrook Dr., San Diego, CA 92111)

Mode-coupling phenomena have long been recognized as contributing to the detailed frequency-dependent behavior of electroacoustic materials and structures. Coupled modes have the effect of repelling each other from the positions in frequency space that they would occupy if the coupling interaction were not operative. Today, mode-coupling effects are most often addressed using finite-element analysis methods. This paper reinvestigates the Poisson (modes coupled via shared inertial mass) mode-coupling issue for simple electroacoustic structures from another perspective. Beginning with the piezoelectric equations of state, equivalent-circuit models are developed. These are useful for interpreting standard forms of impedance/admittance spectroscopy data for a variety of practical experimental geometries. The models are applied to measurement data and their usefulness for performance simulations is illustrated.

3:00

2pEA8. Bragg scattering by subsurface cracks and grooves. Eugene J. Danicki (Polish Acad. of Sci. (IPPT), 21 Swietokrzyska Str., Warsaw, 00-049 Poland)

Bragg scattering of bulk waves from shallow periodic grooves has been proven experimentally to be an effective method for generation of Rayleigh waves in solids. To study this phenomenon, a double layer of alternately positioned periodic cracks is investigated. The system looks like two aligned half-spaces with periodic grooves: cracks correspond to the main grooves' faces; the mechanical connection of the grooves' side faces makes the difference between these systems. A null-space expansion of the wave field is proposed that minimizes the resulting stress. The convenient BIS expansion method is applied in this extensive numerical investigation. For wide cracks, a mode can propagate in the system that resembles the Rayleigh wave because cracks make the substrate surface almost stress free. The efficiency of generation of this mode is studied as a function of crack period and width, the distance between crack layers (corresponding to the groove depth), and the angle of incidence of either

shear or longitudinal waves. The results show that up to 25% of the incident bulk wave power can be transformed into surface waves in an optimal case with appropriate wave beamwidth. [Work supported by PAN-NIST Polish-US MS-Curie Joint Fund.]

TUESDAY AFTERNOON, 5 DECEMBER 2000 SLOOP ROOM, 1:00 TO 3:30 P.M.

Session 2pED

Education in Acoustics: Informal Education in Acoustics

Thomas D. Rossing, Chair
Physics Department, Northern Illinois University, DeKalb, Illinois 60115

Invited Papers

1:00

2pED1. Interactive exhibits illustrating wind instrument acoustics. D. Murray Campbell (Dept. of Phys. and Astron., Univ. of Edinburgh, Edinburgh EH9 3JZ, UK) and Arnold Myers (Edinburgh Univ. Collection of Historic Musical Instruments, Edinburgh EH8 9AH, UK)

The authors were recently involved in developing demonstration equipment for musical instrument museums. The aim of this project, funded by the European Union, was to provide hands-on exhibits which would allow the user to explore some of the basic science underlying the operation of string, wind, and percussion instruments. While it is fairly straightforward to devise exhibits in which the user can pluck a guitar string, strike a tubular bell, or bow a violin, considerations of hygiene and the requirement for skilled embouchure techniques make it impossible to expect members of the public to blow demonstration wind instruments. To overcome this limitation, exhibits have been constructed in which a clarinet and a trumpet are played using artificial mouths. The user can adjust a valve controlling the flow of air from a blower to the artificial mouth; a digital manometer monitors the mouth pressure. The note played can be changed by operating a piston on the trumpet, or a key on the clarinet. The signal from a microphone registering the sound output is fed to the sound card of a PC, and specially written software allows the user to examine the waveform and the frequency spectrum of the sound.

1:30

2pED2. Music and science: Workshops for teachers in elementary schools. Uwe Hansen (Dept. of Phys., Indiana State Univ., Terre Haute, IN 47809)

A number of workshops have been conducted with support from ASA Technical Initiative funds. The emphasis in the workshops is on assisting teachers to use music as a vehicle to introduce science concepts. A monochord is constructed. Teachers are provided a tuning fork, a resonance tube, a spring, a ''singing'' tube, and material to construct a monochord. Concepts discussed include wave properties such as wavelength, frequency, amplitude, phase, and speed of propagation. In addition the connection to music is made by observing standing waves and relating frequency ratios to musical intervals. The teaching approach and workshop content will be illustrated in the presentation.

2:00

2pED3. Websites, video, and audio recordings for informal education in acoustics. Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115)

It is becoming increasingly easier to study acoustics (or any subject) outside the classroom walls. This is partly due to the ready availability of learning materials on the Internet, and video and audio recordings that lend themselves to home study and informal study. We discuss some of the materials that are now available for informal education in acoustics.

2:30

2pED4. Shaping the experience in Experience Music Project: Striking a balance between entertaining and educating in a popular music museum. James Fricke and Andrea Weatherhead (Experience Music Project, 2901 3rd Ave. #400, Seattle, WA 98121)

After eight years of planning, the Experience Music Project opened in Seattle on June 23, 2000. The institution's press materials state that: EMP invites visitors to participate in a total sensory experience: see rare artifacts and memorabilia, hear musicians tell their stories, or play an instrument and create your own music. Learning about music has never been so much fun. Does the experience live up to the hype? EMP's opening exhibitions, activities, and public programs are described, public and critical reception is reviewed, and lessons learned in the first months of operation are shared. The presentation will focus on multimedia exhibit components and interactive exhibits, with special emphasis on the design, development, and objectives of the Sound Lab area, EMP's hands-on activity and learning center.

3:00

2pED5. The CSULB Mobile Science Museum: An interactive science outreach tool. James N. McKibben (College of Natural Sci. and Mathematics, California State Univ. at Long Beach, 1250 Bellflower Blvd., Long Beach, CA 90840, mckibben@csulb.edu) and R. Dean Ayers (California State Univ. at Long Beach, Long Beach, CA 90840)

Since 1980, a converted 27-ft recreational vehicle has been visiting local schools filled with a variety of hands-on science exhibits and activities. The Mobile Science Museum (MSM) has recently undergone a major renovation, including adding a generator. The MSM is now self-contained with inside exhibits that focus on the physical sciences, including sound, light, electricity, etc. The MSM also carries five large tables for outside activities, including a seawater touch-tank containing local intertidal animals, optical illusions, fossils, bones, etc. The majority of the exhibits have been developed or adapted for the MSM and utilize oscilloscopes, oscillators, strobes, power supplies, fans, and other equipment. University students act as docents and serve as enthusiastic role models for our young visitors. It is a compliment to the interest generated by the MSM that many of the current volunteer docents received their first exposure to the program when the MSM visited their schools several years ago. The major funding for the program has come from the CSULB Alumni Association, the Long Beach Industry–Education Council, and other individual support. The requests for visits greatly exceed the current ability to meet them, but this will improve as additional funding becomes available.

Mobile Science Museum Visit

In conjunction with this session, the CSULB Mobile Science Museum will be on display Tuesday afternoon from 1:00 p.m. to 6:00 p.m. in the parking area of the hotel (near the North Tower).

TUESDAY AFTERNOON, 5 DECEMBER 2000 CATAMARAN ROOM, 1:15 TO 4:10 P.M.

Session 2pNSa

Noise, NOISE-CON and Architectural Acoustics: Environmental Noise Focused on Combined Noise Sources

Paul D. Schomer, Cochair
U.S. Army Construction Engineering Research Laboratory, P.O. Box 9005, Champaign, Illinois 61826-9005

Brigitte Schulte-Fortkamp, Cochair
Department of Physics/Acoustics, Oldenburg University, Oldenburg D-26111, Germany

Chair's Introduction—1:15

Invited Papers

1:20

2pNSa1. How many meta-analyses fit on the head of a pin? Sanford Fidell (BBN Technologies, 21128 Vanowen St., Canoga Park, CA 91303)

Meta-analyses of social survey findings on source-specificity of annoyance indicate that the prevalence of annoyance attributable to aircraft noise may be somewhat greater than the prevalence of annoyance attributable to noise produced by surface traffic of similar cumulative level. Likewise, the prevalence of annoyance induced by impulsive noise may differ from that of general transportation noise of similar level. But how annoyed are people who live next to highways that happen to be located near runway ends by the combination of aircraft and surface traffic noise? How annoyed are people who live next to railroad tracks that pass by shooting ranges? One approach to modeling the total annoyance engendered by exposure to multiple noise sources has been to assume that the interactions among source-specific annoyances are negligible, so that total annoyance can be expressed as some form of summation of individual source-specific annoyances. This simplifying assumption may not always be warranted, nor is it often the case that social and physical measurements yield data precise enough to reliably distinguish among alternate summation hypotheses. Policy-related interpretations of the findings of studies of combined annoyance of noises of differing origin should thus be viewed with a healthy skepticism.

1:40

2pNSa2. The acceptor principle in German regulations for environmental noise. Dieter Gottlob and Jens Ortscheid (Federal Environ. Agency, Bismarckplatz 1, 14193 Berlin, Germany)

According to the acceptor principle, environmental exposure shall be assessed from the point of view of the affected persons, i.e., exposures from various sources shall be evaluated as a whole. While the acceptor principle is well established in environmental fields such as air pollution, it is rarely applied in noise regulations. In general, the noise sources are assessed separately. Sometimes the existence of residual noise leads to reduced requirements for the specific noise source under consideration. Important obstacles against

the practical application of the acceptor principle in some countries are inconsistent noise descriptors, measurement and prediction procedures for different sources, and different legal approaches to control noise. Furthermore, there are not enough social field studies available which deal systematically with the assessment of the exposure of several noise sources. Therefore, the proposed procedures for assessing situations with combined noise sources are still lacking scientific support, especially concerning the temporal and spatial distribution of noise sources. In this lecture, the current state of the enforcement of the acceptor principle in German noise regulations will be presented. Some discussions within the framework of the new European noise policy will be reported.

2:00

2pNSa3. Assessment of annoyance in a multisource environment. Truls Gjestland (SINTEF Telecom & Informatics, Trondheim, Norway) and H. M. E. Miedema (TNO, Leiden, The Netherlands)

The annoyance caused by common community noise sources (road traffic noise, rail noise, aircraft noise, etc.) has been shown to be highly correlated with the energy-equivalent noise level, LEQ (or similar derivatives) from that source. In some cases LEQ is adjusted depending on the time of day (as in LDN) or depending on the nature of the source (e.g., impulsive noise). So-called dose-response curves showing the relationship between the noise level and the annoyance have been established for a number of different noise sources. This annoyance assessment method, however, is only valid for single noise sources. The authors present a method for assessment of annoyance in the presence of several noise sources. Each noise source is transformed into a common reference source by means of established dose-response curves, and then added on an energy basis.

2:20

2pNSa4. The importance of loudness functions in assessing noise annoyance. Paul Schomer (USA-CERL-CNN, P.O. Box 9005, Champaign, IL 61826-9005, schomer@uiuc.edu)

Outdoors, Schomer *et al.* have shown that loudness weighting according to ISO 226 can be used to great advantage as a direct substitute for A-weighting. That is, loudness-weighted sound exposure level (LLSEL) can be used as a direct substitute for A-weighted sound exposure level (ASEL). LLSEL, computed with a time constant commensurate with the time constant of the human hearing process, provides significantly better correlation with community response across different noise sources than does ASEL. Indoors, room noise criteria (RNC) curves have been developed using these same human hearing function considerations. The RNC curves and associated computational procedure successfully bridge the gap between the PNC curves of Beranek and the RC curves of Blazier. Taken together, this body of outdoor and indoor research results shows the great importance of the human hearing function in assessing noise annoyance.

2:40–3:00 Break

3:00

2pNSa5. How to define synergetic effects of combined noises. Brigitte Schulte-Fortkamp (Dept. of Phys./Acoust., Oldenburg Univ., 26111 Oldenburg, Germany, brigitte@aku.physik.uni-oldenburg.de)

Evaluation of noise annoyance has to take into account the multidimensionality in a given social situation. Noise annoyance research needs an interdisciplinary procedure methodologically including acoustics, physics, psychology, and sociology. People's judgments depend on multiple factors like environmental area, noise source, characteristic of noises, number of noise events over a day, subjective experience with noises, the social situation, etc. Judging annoyance includes a wide range of variability, which is highly correlated with the variability of living conditions. Considering those contextual conditions of noise annoyance judgments is one important procedure regarding the measurement on annoyance toward noise of combined sources. Taking into account the knowledge on annoyance judgments relying on different subject-oriented conditions, a measurement is needed which refers to objective and subjective parameters. The combination of noise sources is for the judgment of annoyance as well as important subjective parameters. Moreover, the relationship of both defines the background for assessments. Based on an analysis of interviews concerning noises from different sources and on the literature [e.g., Fields *et al.* (1998), Hatfield and Job (1998)], a proposal for an approach in noise annoyance measurement will be presented.

3:20

2pNSa6. What can classical and fuzzy concepts contribute to the analysis of masking effects in environmental noise surveys? Dick Botteldooren, Andy Verkeyn (Dept. of Information Technol., Ghent Univ., St. Pietersnieuwstraat 41, 9000 Gent, Belgium, botteldo@intec.rug.ac.be), and Peter Lercher (Univ. of Innsbruck, Sonnenburgstrasse 16/I, A-6020 Innsbruck, Austria)

In laboratory research masking of one noise source by another is well studied. Field research based on noise surveys fails to show this effect clearly. Both the uncertainty in noise levels and other environmental variables and the fuzziness in annoyance contribute to the difficulty of the problem. Traditional statistical and fuzzy techniques were used to look for masking effects in a large environmental noise dataset gathered in Austria. The fuzzy analyses use a rule-based model that is constructed to predict noise annoyance. Specific attention is paid to rules that predict masking based on variables such as background noise level, distance to masker, direction of bedroom and living room windows, etc.

3:40

2pNSa7. Development of line-based noise pollution reduction programs. Andrew Wearne, Richard Heggie (Richard Heggie Assoc., Level 2, 2 Lincoln St., Lane Cove, NSW, Australia 2066), Suzanne Stuart, and Donald MacLeod (Maunsell McIntyre, Milsons Point, NSW, Australia 2061)

This paper details the results of a unique Noise Pollution Reduction Program sponsored by the Rail Access Corporation of NSW, Australia. The program covers almost 200 km of rail line, located mostly within urban areas of Sydney. The initial stages of the project involved establishing a database of land sensitivities along the selected rail lines, defining all relevant rail noise sources and their characteristics and developing a suite of available noise abatement options and associated costs. The second stage of the project focused on detailed computer noise modeling of the rail lines, coupled with a site-specific field verification study involving all train types and other traffic variables. The third phase of the study involved prioritizing each line and different sections within each line through the development of a unique Community Noise Burden (CNB) index. This index involves assigning relative weightings per given section of line according to the noise components, number and type of dwellings affected, noise exceedance levels, etc. Following the priority ranking of each section of line, a Noise Pollution Reduction Program was then designed, incorporating consideration of community annoyance, cost per CNB improvement, practicality, safety and other benefits/disbenefits.

3:55

2pNSa8. A statistical method for the evaluation of discrete noise sources submerged in background. Alois Heiss (Bavarian State Ministry for Environ. Affairs, Munich, Germany) and Robin T. Harrison (Murphy & Harrison, Inc., Claremont, CA 91711, AirBearZln@aol.com)

Sometimes it is necessary to assess the contribution of a discrete source to the total soundscape even when the overall level of the discrete source is submerged in the background noise. An example of such a situation is the quantification of annoyance from the noise of a factory in the presence of the suburban background of traffic noise, children, birds, etc. The traditional method of making such an assessment, computation and scaling of signal-to-noise ratios in 1/3-octave bands, is not available if spectra of both the discrete source and the background cannot be obtained. Also, this method requires the use of sophisticated instrumentation and time-consuming measurements. This paper describes a methodology which extracts adequate statistical information from stochastic signals to determine the confidence intervals of percentile levels. The Leq of a discrete source in a noisy background can be determined with known high resolution from easily obtained field measurements. The procedure does not depend on frequency analysis. The method is currently being expanded to enable evaluation of level distribution of the source. Basic theory that supports the method is presented and software that facilitates its use is described. An example of its application is given.

TUESDAY AFTERNOON, 5 DECEMBER 2000 SCHOONER ROOM, 1:15 TO 4:10 P.M.

Session 2pNSb

Noise and NOISE-CON: Construction Machinery Noise Limits and Means of Compliance

Daniel R. Raichel, Chair
532 Spencer Drive, Wykoff, New Jersey 07981

Chair's Introduction—1:15

Invited Papers

1:20

2pNSb1. Recent development in European legislation concerning construction equipment. Volker K. P. Irmer (Umweltbundesamt, 14193 Berlin, Bismarckplatz 1, Germany, volker.irmer@uba.de)

European noise legislation since the 1970s consisted of directives on the approximation of the laws of the Member States in order to ensure the smooth functioning of the internal European market. In the late 1990s, the protection of human health and well-being started to play a more important role. A new EU directive on noise emission from equipment for use outdoors replacing former directives concerning construction machines is in force since its publication in the Official Journal of the European Communities on 3 July 2000. It lays down provisions concerning the marking of equipment with its guaranteed sound power level for about 60 types of outdoor equipment (construction machinery, hobby and garden equipment, and some others) and for about 20 of those noise limits in two steps. Noise limits are demanding but the current state of the art of noise reduction allows compliance without excessive additional costs. The current directive states that only equipment complying with these provisions may be placed on the European market that consists of 375 million citizens. The European Directive now must be transposed in national legislation; the provisions have to be applied from 3 January 2002.

1:40

2pNSb2. Railway construction in Hong Kong: How to make it work. Glenn H. Frommer (Mass Transit Railway Corp., MTR Tower, Kowloon Bay, Kowloon, Hong Kong SAR, PROC)

The Mass Transit Railway Corporation designs, constructs, and operates one of the world's most intensively used railway networks with 2.3 million passengers daily on 70 km of track. New railway construction is often undertaken in densely populated urban environments with high-rise developments surrounding works sites. For the Lantau and Airport Railway, operational in 1998, the Corporation had to remove 1000 m^3 of granite 7 m from a high-rise residential estate for the construction of Lai King Station. A

lightweight noise enclosure was erected but not without significant noise and political impacts. However, the station was successfully completed on time and within budget. Current extensive modifications to Mei Foo Station 15 m from residential towers require a heavyweight noise enclosure to include vehicular emergency service access. The construction method was developed to minimize noise impacts. The skeleton of the enclosure is now in place and there have been no significant environmental problems. This presentation will detail the following: the analysis methods used in determining noise impacts; the structure and training employed within contracts to obtain compliance to the noise specification; and the manner in which the noise impacts are monitored and controlled.

2:00

2pNSb3. Noise control for a vacuum excavator: A mutually beneficial solution. Erich Thalheimer (Central Artery/Tunnel Project, 185 Kneeland St., Boston, MA 02111, exthalhe@bigdig.com)

The Central Artery/Tunnel (CA/T) Project in Boston (the Big Dig) makes use of several vacuum excavator trucks (vac-trucks) in order to reduce accidental utility disruptions. Unfortunately, these vac-trucks generate very loud low-frequency noise due to their vacuum blowers which can cause nearby buildings to vibrate and has led to significant numbers of complaints from distressed abutters. In reaction, the CA/T Project undertook an intensive study to determine what mechanisms were producing the offensive low-frequency noise, and to devise mitigation measures to effectively control the low-frequency noise so as to allow the vac-trucks to continue working without causing noise complaints. Working cooperatively with the manufacturer (Guzzler) as well as with the contractor (Walton Systems), new and improved (Universal) silencers were retro-fitted to a test vac-truck and the improvements in low-frequency insertion loss were measured and documented. The new silencers effectively reduced the 40–125-Hz third-octave band range noise by 10–15 dB, thus approaching the noise control design goals selected to avoid causing nearby buildings to vibrate. This is a prime example of how cooperative efforts between vehicle manufacturers, construction contractors, and project officials can lead to mutually beneficial results.

2:20

2pNSb4. Noise monitoring, control, and compliance practices at several construction sites in Southern California. Douglas K. Eilar (Douglas Eilar & Assoc., 321 N. Willowspring Dr., Encinitas, CA 92024)

Noise protection for residential neighborhoods, commercial properties, and wildlife habitats near construction sites is achieved to varying degrees within a maze of local, state, and federal regulations. Standards for a given site typically include noise limits for public nuisance under the municipal ordinance, but a project may also be subject to more restrictive requirements emanating from conditions of approval from a discretionary permit and the environmental review process. Sometimes, goals or provisions of the municipality's noise element to the general plan are a factor. Sometimes, only noise monitoring is required, for documentation or to determine if noise is a problem. For wildlife, federal or state agencies may have established limits applicable to habitats of certain sensitive, endangered, or threatened species, during nesting season. Noise limits for constructing a subway access in Los Angeles require measurement of maximum noise level (Lmax), while noise limits for grading a park in San Diego use hourly average sound level (Leq). Different limits apply to intermittent versus continuous noise, to different times of day, and to different land uses. Noise reduction may be attained by means such as sound-attenuation barriers, certification of equipment noise emission prior to use, or operational controls.

2:40–3:00 Break

3:00

2pNSb5. Survey of worldwide legislation on the control of construction noise. Kai Ming Li (Dept. of Mech. Eng., Hong Kong Polytechnic Univ., Hung Hom, Hong Kong)

The sustained economic growth in many countries of the world in the past few years has led to significant increases in the construction activities for building new homes, new infrastructures such as highways, and airports for cities. This construction work inevitably causes an increase in noise levels. Indeed, the problem of noise is one of the most important environmental issues in recent years. With increased exposure and the greater sensitivity of much of the population, the number of noise-related complaints has grown at an alarming rate over the past few years. Current legislation imposes a limit of acceptable noise levels, but seems ready to be tightened further in reflection of popular demand. This paper reports a recent survey on the worldwide legislation on the control of construction noise. The survey concentrates on different countries in the Pacific Rim, the USA, and a few countries in the European Union. Relevant guidelines proposed by such international organizations as the World Health Organization and the Organization for Economic Cooperation and Development are also reported.

3:20

2pNSb6. Legislation concerning construction noise in Germany. Eberhard Fischer-Sheikh Ali and Volker K. P. Irmer (Umweltbundesamt, Bismarckplatz 1, 14193 Berlin, Germany)

German legislation concerning construction noise started in 1965. It covers noise exposure in the neighborhood due to construction sites as well as noise emission from construction equipment. Noise exposure caused by construction sites may not exceed values specified in an administrative provision. These values differ for various types of residential areas and for different periods of the day. If they are, however, exceeded, the administration shall take the appropriate measures to limit the exposure. Noise emission from certain construction equipment is regulated by an ordinance; it contains noise limits that must be observed by the manufacturers placing equipment on the market and the obligation to label all equipment is specified with the sound power level guaranteed by the manufacturer. The limits depend on the type of equipment and its installed power. The Environmental Label (Blue Angel) is a voluntary product designation, it is one of the so-called soft instruments of environmental policy. The Blue Angel for construction equipment is awarded to machines that yield a noise emission 5–12 dB(A) less than noise limits of the ordinance mentioned above.

3:40

2pNSb7. Noise and vibration control for the Los Angeles Metro Rail Construction Project. Thanh Luc (Parsons Engineering Science, 100 W. Walnut St., Pasadena, CA 91124)

Parsons Engineering Science has been retained by the Los Angeles Metropolitan Transportation Authority as the noise and vibration consultant for construction of the 17-mile-long Metro Rail subway system for a period of more than 10 years. Throughout various construction phases, short-term and long-term noise and vibration monitoring have been conducted at sensitive receptor sites such as residences, schools, hospitals, recording studios, businesses, historic buildings, and hotels, to measure both background and construction activities. Continuous noise and vibration data are collected remotely over a computer terminal, such that levels can be monitored on a moment's notice; our engineers can hear the actual sounds at the remote monitoring site from the office (15–20 miles away), and simultaneously receive decibel levels reported in real-time. Studies are conducted to determine the impacts of various construction activities, such as pile installation, and analyze the feasibility of various construction methodologies based on project noise and vibration compatibility criteria. Mitigation measures are developed to reduce impacts. Examples of mitigation measures included the use of noise curtain walls, noise equipment specifications, selection of construction methodologies, and various acoustical treatments for noise sources. Through the monitoring program and numerous studies, noise and vibration impacts have been successfully minimized.

3:55

2pNSb8. Demonstration hearing conservation program for construction workers in Washington State. Alice H. Suter (Alice Suter and Assoc., 575 Dogwood Way, Ashland, OR 97520)

Construction workers have traditionally been an ''underserved'' population in the area of noise control and hearing conservation. Approximately 500,000 of these workers are currently exposed to potentially hazardous levels of noise and little progress has been made to the U.S. to mitigate this exposure. A labor–management–professional team composed of members of the Washington State Building and Construction Trades Council, the Building Trades Labor Management Organization, the Center to Protect Workers' Rights, and professionals in industrial hygiene, audiology, and medicine, have developed a ''common sense'' hearing conservation program to serve construction workers and employers. Initiated as a demonstration program in the State of Washington, the program is tailored to the U.S. construction industry. It will include worker and supervisor training using construction-specific materials, training in the fundamentals of construction noise control, fitting and training in the use of hearing protection devices, audiometric surveillance, centralized record keeping, and the evaluation of program effectiveness. Special emphasis is being placed on the use of readily available noise control solutions as well as hearing protection devices that will not degrade communication and the perception of warning signals. Program plans are currently being finalized and training will be initiated in December of 2000.

TUESDAY AFTERNOON, 5 DECEMBER 2000 BRIGANTINE ROOM, 1:20 TO 3:40 P.M.

Session 2pNSc

Noise and NOISE-CON: Vehicle Noise and Vehicle Noise Tests

Paul R. Donavan, Chair
General Motors Corporation, N & V Center, 8916 Cooley Lake Road, White Lake, Michigan 48386-4027

Contributed Papers

1:20

2pNSc1. Rolling noise ANC system: An experimental facility test. Federico Rossi (Laboratorio di Acustica, Universit di Perugia, Via G. Duranti 1/A-4, 06125 Perugia, Italy)

A new active system for reduction of rolling noise produced at tire-ground contact is proposed. The system is constituted by acoustical emitters installed on the car mudguard close to tire-ground point of contact. Active control is achieved by means of an electronic control unit which generates the control signal according to a digital error minimization technique. The active system has been tested on a stationary condition by means of an original experimental facility built at Acoustic Laboratory of the University of Perugia. The facility reproduces the acoustical emission of rolling noise employing a recorded noise source. Experimental tests show that the proposed active noise control system is particularly suitable for rolling noise low-frequency components attenuation. Application on a traveling car is actually going on.

1:35

2pNSc2. High-temperature quiet-flow test facility. Jim R. Cummins, Jr., Jim B. Causey, and David Kapsos (Universal Silencer, P.O. Box 411, Stoughton, WI 53589, US@Universal-Silencer.com)

A high-temperature, quiet-flow test facility for the testing and evaluation of exhaust silencers and silencer materials has been developed. This facility is driven by a rotary positive displacement blower delivering up to 110 m^3/min (4000 cfm) of airflow. This airflow is put through both a passive and active silencer to achieve minimal uncontrolled noise in the test stream. Controlled sound is injected into the airflow by a set of loudspeakers in a mixing/settling chamber. The airflow is heated up to 540 °C [1000 °F] by a 144-kW electric heater just prior to entering the test section. The test section is rated for a 10-cm- (4 in.) diam flow path and up to 1.8-m- (6 ft) long sample. The outlet of the test section exhausts outside the test building and a microphone measurement system measures the radiated sound power level. This facility allows tests to be run with flow, sound level and temperature controls to obtain *in situ* measurements of silencer performance including insertion loss.

1:50

2pNSc3. Sound transmission characteristics of bulb seals. Junhong Park, Luc G. Mongeau, and Thomas Siegmund (School of Mech. Eng., Purdue Univ., 1077 Herrick Labs., West Lafayette, IN 47907-1077)

Sound transmission through door and window sealing systems is one important contributor to interior noise in vehicles. The noise generation mechanism involves the vibration of the seal due to the unsteady wall pressures associated with turbulent flow over the vehicle. Bulb seals are often used in cars, an application where the seals experience large deflection variations due to build variation between vehicles. The performance of a bulb as a sound barrier was the focus of this study. The vibration response of a bulb seal together with the radiated sound was modeled using a finite element method. Material properties for the elastomeric foams were measured and used in the FEM. Acoustic-structure interactions were considered and the noise generated on the quiescent side of the

seal was calculated. Experimental data was obtained for several bulb seal specimens using a reverberant room test method. The radiated sound intensity was measured using an intensity probe. The seal wall vibration was measured using a laser vibrometer. The transmission loss of the bulb seal was calculated from the measured data. It was compared with the value from the numerical simulations. The effects of bulb seal design parameters such as density, elastic modulus, and damping ratio were investigated.

2:05

2pNSc4. Vehicle door seal modeler for wind noise. Yuksel Gur (Ford Res. Lab., P.O. Box 2053/MD3135-SRL, Dearborn, MI 48121, ygur@mail.ford.com)

Wind-noise performance of a passenger vehicle plays a significant role in the customer's perception of overall vehicle quality. Wind noise aerodynamically generated on the exterior surface of a vehicle can propagate to the vehicle interior through body panels, window glass, and the seal gap between the door and door opening panel. Experimental studies indicate that automotive door seal systems are generally a major contributor to the vehicle wind-noise performance. Traditionally, automotive door seals are designed using the ''bread board'' and other ''build and test'' procedures. Two new finite-element-based acoustic-structure analysis tools have been developed to evaluate the wind-noise performance of door sealing system. These validated CAE techniques allow the engineer to predict door seal system performance in sound transmission and aspiration leakage before the availability of prototype hardware. These methods combine the separate analysis capabilities of finite-element analysis for acoustics, rubber elasticity, steady-state harmonic viscoelasticity, and a unique modeling approach to perform a fully coupled structure/acoustic interaction analysis of door seal systems. These unique CAE methods, experimental validations, and implementation of these techniques on product design will be presented.

2:20–2:35 Break

2:35

2pNSc5. Maglev—A super fast train. Areg Gharabegian (Parsons Engineering Science, Inc., 100 W. Walnut St., Pasadena, CA 92114, Areg@gharabegian@parsons.com)

Magnetic levitation (Maglev) is a technology where magnetic forces lift, propel, and guide a vehicle over a guideway. This configuration eliminates contact between vehicle and guideway and permits speeds of up to 300 mph. The Federal Railroad Administration is evaluating the feasibility of Maglev construction in seven cities. Parsons Engineering Science conducted noise and vibration studies for four of these cities. Use of magnetic levitation and propulsion minimizes noise from mechanical and moving parts; however, the benefit of low-noise emission diminishes as a vehicle operates at high speed. Aeroacoustic sources generally dominate noise levels at speeds of 120 mph or greater. High-speed Maglev passby is characterized by high-noise levels and brief duration, which may startle people who are close to the alignment. Maglev systems generate groundborne vibration when a vehicle travels above the guideway due to the sudden on and off load of the vehicle and magnetic current. Maglev noise, vibration, and startling sources as well as impacts and mitigation measures will be presented. Available vibration data are based on German standards because there is a prototype Maglev train in operation in Germany. These standards and procedures that are used to calculate equivalent vibration levels in in./s will be presented.

2:50

2pNSc6. Relation between noise emission, friction and rolling resistance of car tires—Summary of final results. Ulf S. I. Sandberg (Swedish Natl. Rd. and Transport Res. Inst., SE-58195 Linkoping, Sweden, ulf.sandberg@vti.se) and Jerzy A. Ejsmont (Tech. Univ. of Gdansk, PL-80952 Gdansk, Poland)

A European Union directive with noise emission limits for noise from vehicle tires is underway. It has been suggested that tires should meet certain limits also with regard to skid resistance, assuming a conflict between requirements for low noise and high friction. It is important to study whether there is any conflict between requirements with respect to noise emission from tires and other important performance characteristics. Therefore, the authors conducted an extensive experimental study on this topic in 1997–1999. Tire/road noise was measured with the CPX method on three road surfaces (8 mm medium-textured SMA, 16 mm smooth-textured DAC, and 12 mm rough-textured surface dressing) at two speeds. Noise was also measured with the drum method on three drum surfaces (replica ISO10844 surface, smooth-textured surface and rough-textured surface). Friction was measured with friction tester BV12, using optimum slip and locked wheel, at 70 km/h, on a smooth-textured DAC. Rolling resistance was measured on a laboratory drum facility, using a smooth safety walk surface and a rough-textured surface. Correlations, for 75 to 100 car tires, between all measured variables have been calculated. It is shown that there is no significant conflict between low noise and high friction, nor between low noise and low rolling resistance.

3:05

2pNSc7. The effect of damping of the disc brake squeal noise. Kihong Shin, Jae-Eung Oh, Wootaek Kim (Dept. of Automotive Eng., Hanyang Univ., 17 Hangdang-Dong, Seongdong-Gu, Seoul 133-791, South Korea), and Mike Brennan (Univ. of Southampton, Southampton SO17 1BJ, England)

A novel mathematical two-degree-of-freedom model is introduced to describe the dynamics of a disc brake system. The model is emphasized on the interacting dynamics between disc and pad components, and is investigated to understand the mechanisms of the disc brake squeal noise. Sensitivity analysis is studied to find unstable regions considered as the squeal state. Time signals are also collected from a set of first-order differential equations of the model. The signals are then used for phase space analysis to investigate the size of limit cycles. Both sensitivity analysis and phase space analysis are focused on the effect of damping of the system. The results show that damping treatment should be done on both disc and pad, whereas damping treatment has been done on the pad only in most practical situations.

3:20

2pNSc8. An auditory model for the prediction of detection thresholds of impulsive sound events. Scott Amman, Mike Blommer, and Jeff Greenberg (Ford Motor Co., Sci. Res. Lab., 20000 Rotunda Dr., P.O. Box 2053, MD 2115, RM 3135, Dearborn, MI 48121)

Impulsive sound events such as squeaks and rattles experienced in a vehicle can be a major source of customer dissatisfaction. These events almost always occur in the presence of some sort of background noise (e.g., wind, road and powertrain noise). When complete elimination of the impulsive sound event is either not possible, or too costly, the same effect can be had by pushing the level of the sound below the detection threshold. This paper describes the development of an auditory model that has the ability to predict detection thresholds of impulsive sound events in the presence of noise. The model was initially developed using equal-energy exponentially damped sinusoids with center frequencies of 250, 500, 1000, 2000, 4000 and 8000 Hz mixed with pink noise. It was then validated using four different squeak and rattle sounds each mixed with three background noises recorded in a vehicle (wind, smooth and rough road noise). An up–down Levitt procedure was used for threshold determination. Two subjects participated in the first experiment and six in the second. In general, the developed model was able to predict detection thresholds with an error that is on par with the subject-to-subject variability observed in the experiment.

Session 2pNSd

Noise and NOISE-CON: Power Plant Acoustics Round Table

Robert A. Putnam, Chair
Siemens Westinghouse PC, MC 590, 4400 Alafaya Trail, Orlando, Florida 32826

A panel of invited speakers from Session 2aNSb: Power Plant Noise Control and Prediction and Industrial Noise will discuss issues confronting the accurate determination of emissions and immissions from power plants. Panelists will introduce issues including prediction, noise control techniques, and measurements and encourage questions and comments from the audience.

Session 2pSA

Structural Acoustics and Vibration, Signal Processing in Acoustics and Physical Acoustics: Diagnostics of Vibration and Noise in Structures II

Scott D. Sommerfeldt, Chair
Department of Physics and Astronomy, Brigham Young University, Provo, Utah 84602-4673

Invited Papers

1:30

2pSA1. Meanings of SVD and wave-vector filtering in the near-field acoustical holography using the inverse BEM. Jeong-Guon Ih (Dept. of Mech. Eng., Korea Adv. Inst. of Sci. and Technol., Science Town, Taejon 305-701, Korea) and Sung-Chon Kang (Daewoo Heavy Ind. Ltd., 6 Manseok-dong, Dong-ku, Inchon 401-010, Korea)

In the conformal near-field acoustical holography (NAH), using the boundary element method (BEM), the vibroacoustic information on the source surface can be indirectly reconstructed by utilizing the measured field pressure and the inverse transfer matrix. The involved vibroacoustic transfer matrix is generally ill conditioned and the reconstruction process should include the singular value decomposition (SVD) in order to solve the problem in the inverse procedure. The accuracy of the reconstructed field is deteriorated substantially due to the ill conditioning of the transfer matrix and the inevitable measurement noise of field pressure related to the nonpropagating wave components. In this study, the computational processes of SVD and wave-vector filtering in the BEM-based NAH are discussed and their physical meanings are investigated through a simulation example. The vibroacoustic transfer matrix decomposed by SVD permits an effective regularization of the source field: This is possible because the involved information in the transfer matrix can be separated into the radiation efficiency and the wave modes at the source and surface fields. In particular, it is clearly shown that the restored source image can be improved dramatically by adopting the wave-vector filter that suppresses the nonpropagating wave components appropriately that cause the deterioration of the reconstruction accuracy.

2:00

2pSA2. Moving frame acoustic holography and its possible applications in machine monitoring and diagnostics. Yang-Hann Kim and Soon-Hong Park (Ctr. for Noise and Vib. Control, Dept. of Mech. Eng., KAIST, Science Town, Taejon 305-701, Korea, yhkim@mail.kaist.ac.kr)

Moving frame acoustic holography has been successfully applied to predict sound pressure distribution on the surface of interest to us. It has been motivated to enlarge the aperture size of the hologram. In the meantime, however, a great deal of application has been found in what is related to making sound image with regard to time as well as frequency. The method is, in fact, analogous to a copy machine. We make the image of sound, instead of duplicating light image, which has to do with a copy machine. The theory [Y.-H. Park and S.-H. Kim, J. Acoust. Soc. Am. **104**, 3179–3189 (1998)] shows that we can travel an array of microphones over the surface of which we want to have a sound/noise image. The theory also permits one to recover the sound image from the data obtained by the array microphone fixed on the ground. The sound image is the spatial characteristics of sound that will be observed by the microphone system attached to the sources. This paper reviews possible areas of applications for a train or subway, or a quality control system in a factory.

2:30

2pSA3. Characterizing mechanical system integrity using structural surface intensity. Jeffrey C. Banks and Stephen A. Hambric (217 Appl. Sci. Bldg., Penn State Univ., University Park, PA 16802, jcb242@psu.edu)

Many structural intensity measurements have been made on simple beam and plate-like structures to determine source strengths and transmission paths. When analyzing more complex structures it is often difficult to compute the intensity throughout the depth of an element. Using a method developed by Pavic [G. Pavic, J. Sound Vib. **115**, 405–422 (1987)], the structural surface intensity (SSI) is computed on a gearbox housing using an array of accelerometers and strain gauges. The focus of this research is to use SSI as a machinery fault indicator, with the objective of developing a method that is more sensitive than other conventional diagnostic methods. Experimental data were taken on a gearbox from the Mechanical Diagnostic Test Bed (MDTB) at the Penn State Applied Research Laboratory as the gearbox was run to failure. The changes in energy flow are analyzed and characterized to provide criteria for fault detection. Analysis of the data indicates that significant changes in structural surface intensity occur as geartooth faults develop and that SSI can be more sensitive than other traditional indicators. [This research was supported by the Multidisciplinary University Research Initiative for Integrated Predictive Diagnostics (Grant No. N00014-95-1-0461) sponsored by the Office of Naval Research.]

2:45–3:00 Break

3:00

2pSA4. Vibration reduction characteristics of an isolated concrete test support. Hooshang Khosrovani (Veneklasen Assoc., 1711 16th St., Santa Monica, CA 90404, hkhosrovani@veneklasen-assoc.com)

Some testing and research facilities require extremely low vibration background environments. Specially designed, isolated concrete supports are built to provide the required levels of isolation. The vibration isolation characteristics for an isolated concrete support have been measured and results are compared to predicted values.

3:15

2pSA5. Seismic wave attenuation coefficient for soils. Hooshang Khosrovani and Jose Ortega (Veneklasen Assoc., 1711 16th St., Santa Monica, CA 90404, hkhosrovani@veneklasen-assoc.com)

Several projects over the past years have provided the opportunity to measure the ground propagation properties of seismic waves. Data from a variety of vibration sources and soil types have been summarized in the form of the coefficient of attenuation versus frequency in the ranges 5–250 Hz. Also included in the summary are data from the published literature. Although the data exhibit a wide scatter their availability will enhance the effort to predict induced ground vibration. This information may in turn be used in predicting noise and vibration levels in structures.

3:30

2pSA6. Visual engineering environment based on interchangeable virtual instrument for acoustic pressure field. Huancai Lu (Acoust. Soc. Am., Dept. of Instrumentation Sci. & Eng., Zhejiang Univ., Hangzhou 310027, PROC, hclu@moi-lab2.zju.edu.cn)

Architecture and kernel technology of a visual engineering environment for an acoustic pressure field is presented. The system software is constructed based on the hierarchy architecture of virtual instrument software and COM technique in order to realize the flexible reconfiguration of an acoustic pressure field measuring system. The attribute models of measuring channels and sensors are established for instrument's and sensor's interchangeability. The locking critical section technique is used for the security of multitask intercurrent process and share relation database. These kernel techniques make acoustic pressure field measuring system easily upgraded (e.g., the function and scale of the measuring system can be extended for different test tasks), instruments or sensors can be interchanged without modifying the software source codes. Therefore maintenance of the acoustic pressure field measuring system is greatly reduced and the lifecycle is prolonged. In this visual engineering environment, the measuring data of acoustic pressure field can be controlled in real-time and expressed in table, curve, bar chart, contour, and virtual instrument. Simultaneously the data can be transmitted through an ODBC interface to another computational environment (e.g., MATLAB) or Internet client for further analysis.

3:45

2pSA7. Discharge acoustic signal detection in insulating oil using the optical fiber Sagnac interferometer. Jongkil Lee (Mech. Education, Andong Natl. Univ., Andong 760-749, Korea), June-ho Lee (Ho Seo Univ., Korea), and Sang-joon Kim (KEPRI, Korea)

In this paper, an optical fiber sensor (OF sensor) utilizing the principal of the Sagnac interferometer was proposed to detect the discharge acoustic signals which generated from a needle-sphere electrode system in insulating oil. The performance of the OF sensor was checked by a sinusoidal calibration signal generated by a PZT actuator at 198 kHz. The detected discharge signals consisted of the acoustic signal and electrical noise. The noise signal could be removed by a digital low-pass filter. It was demonstrated that the OF sensor in this research had a possibility to detect the discharge acoustic signals in the power apparatus.

4:00

2pSA8. A crack localization system based on the spatio-temporal gradient method. Kenbu Teramoto (Saga Univ., 1-Honjo Saga-shi, 8408502 Japan)

For several decades, nondestructive tests, including x rays, gamma rays, radar, infrared thermography, and acoustic methods have been used not only for concrete structures but for other types of materials. Acoustic methods are the oldest and most widely used nondestructive testing. They are based on the propagation, and in some cases reflection, of elastic waves in solids or over their boundary surfaces. A well-known example is striking an object with a hammer and listening to variations in the ''ringing'' sound to detect the presence of internal voids, external cracks, or other defects. Usually, there are three techniques based on acoustic nondestructive tests sorted into the wave-velocity method, resonance methods, and echo methods. This paper proposes a fourth method with spatio-temporal analysis, which has two processes, as follows. In the observation, the resulting surface wave, its partial derivatives for x, y, and t, are calculated at several points. In the detection, the rank of the short-duration covariance matrix of time-series vector $\mathbf{f}$ $[=(f\ f_x\ f_y\ f_t)^T]$ classifies the situation of the surface-wave field. Consequently, the proposed method can provide the wave front and the location of the cracks as the singularity in the sound field.

4:15

2pSA9. Usage of Wigner–Ville distribution for vibroacoustical forced oscillation diagnostics of cracks. Leonid M. Gelman (Dept. of Nondestructive Testing, Natl. Tech. Univ. of Ukraine, 37, Peremogy Pr., Kiev, 252056, Ukraine) and Denis A. Adamenko (Natl. Tech. Univ. of Ukraine, Kiev, 252056, Ukraine)

For diagnostics of fatigue cracks using vibration and noise, the new low-frequency vibroacoustical method of forced oscillations with nonstationary excitation is considered. The usage of an advanced time-frequency signal-processing technique based on Wigner–Ville distribution is proposed and developed to improve the effectiveness of diagnostics. The simulation results are presented.

4:30

2pSA10. Detection of fatigue cracks by wavelet decomposition of the nonstationary narrow-band acoustical signal. Nadejda I. Bouraou (Dept. of Orientation and Navigation Systems, Natl. Tech. Univ. of Ukraine, 37 Peremogy pr., Kiev, 03056, Ukraine, nadya@burau.inec.kiev.ua) and Pavel I. Marchuk (Natl. Tech. Univ. of Ukraine, Kiev, 03056, Ukraine)

This paper presents new results of the signal processing of the nonstationary narrow-band acoustical signal for vibroacoustical diagnosis of aircraft engine blade fatigue cracks. The diagnosis method uses nonstationary vibratory excitation with a variable central frequency at the engine running start. The above-mentioned excitation is considered to be a linear-frequency-modulated signal with amplitude modulation by a stationary random process. The cracked blade is considered to be piecewise-linear system. Assuming these preconditions, the nonstationary narrow-band acoustical signal, radiated by the blade, is simulated and processed by using the wavelet decomposition technique. Fault detection is carried out at the approximation of the last (fifth) level of decomposition. The influence of the crack size increasing on the considered level approximation is investigated.

TUESDAY AFTERNOON, 5 DECEMBER 2000 NEWPORT ROOM NORTH, 2:00 TO 5:00 P.M.

Session 2pSC

Speech Communication: Measuring and Modeling Speech and Voice (Poster Session)

Bruce R. Gerratt, Chair

Division of Head and Neck Surgery, University of California, Los Angeles, 31-24 Rehab Center, Los Angeles, California 90095-1794

Contributed Papers

All posters will be on display from 2:00 p.m. to 5:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 2:00 p.m. to 3:30 p.m. and contributors of even-numbered papers will be at their posters from 3:30 p.m. to 5:00 p.m.

2pSC1. Effects of oscillation of a mechanical hemilarynx model on mean transglottal pressures and flows. Fariborz Alipour (Dept. of Speech Pathol./Audiol., Univ. of Iowa, 334 WJSHC, Iowa City, IA 52242) and Ronald C. Scherer (Bowling Green State Univ., Bowling Green, OH 43403)

This study introduces a mechanical model of the larynx for investigating the aerodynamics of phonation. The model mimics the hemilarynx. The tracheal inlet section was rectangular (2.5 cm by 2 cm). The vocal fold was fabricated from hard plastic with an attached oscillating plunger. A speaker assembly and audio amplifier drove the plunger, mimicking one-dimensional vocal fold motion toward a flat wall. The glottal shape was rectangular. The glottal diameter was well specified or dynamically followed with a laser system. The air was sucked through the channel using a vacuum with controlled speed. Frequency and amplitude of the glottis were varied. The mean pressure and mean flow data were recorded. For steady flow conditions, the glottal gap ranged from 0.4 to 2.5 mm. The pressure coefficient for steady flow had a range of 3.1–1.3 for Reynolds numbers between 500 and 5000. For oscillation conditions: (a) the frequency was varied from 75 to 150 Hz while the amplitude was held (almost) constant; and (b) the amplitude was varied for a fixed frequency of 100 Hz. The results indicate that oscillation amplitude has a strong effect on the flow resistance of the glottis. [Work supported by NIDCD Grant No. R01 DC03566.]

2pSC2. Airflow patterns in hemilarynx models. Fariborz Alipour and Douglas Montequin (Dept. of Speech Pathol./Audiol., Univ. of Iowa, Iowa City, IA 52242)

Airflow patterns of a hemilarynx were simulated numerically by the method of computational fluid dynamics (CFD) and were investigated experimentally with a flow visualization method. In the numerical model, vocal fold motion was simulated by sinusoidal variations of the inferior and superior glottal widths at different phase conditions. The straight glottal wall alternated between converging and diverging shape, while the input airflow varied sinusoidally between zero and a maximum. The control variables were oscillatory frequency, Reynolds number, and phase difference of the inferior and superior sections used for glottal motion. Flow visualization was performed using a hemilarynx physical model with a rectangular-shaped glottis. Using a Kodak high-speed camera, airflow patterns were recorded for various glottal widths and flow rates. CFD simulation described the complete airflow pattern and velocity distribution which indicated the existence of a high-speed jet near the fixed wall, flow separation, and vortex formations behind the glottis. Flow visualization also revealed similar patterns of flow separation, near-wall jet, and vortices below the jet. [Work supported by NIDCD Grant No. R01 DC03566.]

2pSC3. Direct measurement of the glottal volume velocity waveform in an *in vivo* canine model. Andrew Verneuil, Jody Kreiman, Bruce R. Gerratt (Div. of Head and Neck Surgery, UCLA School of Medicine, 62-132 CHS, Los Angeles, CA 90095, verneuil@ucla.edu), Vijay Dhir (Univ. of California, Los Angeles, CA 90095), and Gerald Berke (UCLA School of Medicine, Los Angeles, CA 90095)

Phonation occurs when the potential pressure energy provided by the lungs is translated into the kinetic energy of a glottal air pulse by the vibrating vocal folds. The primary acoustic source energy is described by the amount of air passing the glottis (volume) and the instantaneous speed at which it is moving (velocity). Indirect methods of acoustic analysis have provided estimates of the volume-velocity waveform. This study provides direct measurements by applying a constant temperature anemometer and videostroboscopic glottic analysis in an *in vivo* canine model. The velocity profile across the glottis and glottal area were measured in the anterior, middle and posterior glottis. These measurements were then temporally synchronized and combined to create a volume-velocity waveform of the acoustic source. Consistent with previous studies, velocity waveforms were double peaked across the glottis and were

usually greatest anteriorly. Significant contributions to the volume-velocity waveform were made by the anterior and middle portions of the vibrating vocal folds, while the posterior glottis had less contribution. Inertial delays and flow reversals were demonstrated during the closed phase of phonation. Understanding the measured volume-velocity waveform provides more insight into the acoustic source spectrum and its role in phonation.

2pSC4. Flow visualization in a model of the glottis with a symmetric and oblique angle. Daoud Shinwari, Ronald C. Scherer (Dept. of Commun. Disord., 200 Health Ctr., Bowling Green State Univ., Bowling Green, OH 43403, ronalds@bgnet.bgsu.edu), Abdollah A. Afjey, and Kenneth DeWitt (Univ. of Toledo, Toledo, OH 43606-3390)

Modeling the human larynx can provide insights into the nature of flow within the glottis. This study reports intraglottal pressure profiles and flow visualization for a symmetric and an oblique glottis with a glottal angle of 10 deg divergence. For the oblique case, the glottis slanted at an angle of 15 deg. A Plexiglas model of the larynx was used. Each vocal fold had at least 11 pressure taps. The minimal glottal diameter was held constant at 0.04 cm. Each case was subjected to steady airflow corresponding to transglottal pressure drops of 3, 5, 10, and 15 cm H_2O. Pressure profile results showed that pressures were different on the two sides of the glottis; these data were strongly supported by an earlier study using a different model. Flow visualization in all cases showed that flow separated from one side of the glottis and remained attached to the other. For the oblique case, the separation point on the divergent wall moved upstream in the glottis with greater flows. The laminar core of the skewed jets decreased in length with higher flows. The jet caused asymmetric circulating regions downstream of the glottis in the reservoir section. [Work supported by NIH.]

2pSC5. Parameters for a first-order Kelvin model of laryngeal muscles. Eric J. Hunter and Ingo Titze (Dept. of Speech Pathol. and Audiol., Natl. Ctr. for Voice and Speech, The Univ. of Iowa, Iowa City, IA 52240, eric-hunter@uiowa.edu)

Because the intrinsic muscles of the larynx determine vocal fold posturing in phonation, a model of vocal fold posturing requires data from the passive and active viscoelastic properties of these intrinsic muscles. Using reported one-dimensional active (twitch and tetanus) and passive muscle (step elongation and relaxation) measures from fresh excised canine laryngeal muscles, one-dimensional first-order Kelvin model parameters were obtained. Obtaining these parameters was accomplished through optimization, where force-elongation patterns predicted by the model were compared to measured muscular passive cyclic force-elongation patterns. These parameters are reported along with the design for the Kelvin model. The results of this study, which will be integrated into a larger finite-element posturing model, become the first step in predicting the vocal fold configurations needed for phonation.

2pSC6. New methods for objective evaluation of nasality. Masanobu Kumada (Dept. of Speech Physiol., Univ. of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033 Japan), Toshiaki Kaneko (Sci. Univ. of Tokyo, Tokyo, 162-0825 Japan), Takahiko Ono, Masanao Ohashi, Koji Ishida (Ono Sokki, Yokohama, 226-8507 Japan), and Seiji Niimi (Univ. of Tokyo, Tokyo, Japan)

New methods for objective evaluation of nasality were introduced, including laser Doppler vibrometer, contact microphone system, and small microphone system. Subjects were two Japanese healthy males who speak Tokyo dialect. Tasks were sustained phonations of five Japanese vowels and /m/, repetitions of /ai/ and /mami/, and Japanese words of /shinksnsen/ and /shinbunshi/. Vibration of nasal skin was detected by laser Doppler vibrometer in terms of velocity (Nv), and by contact microphone in terms of deviation (Nc). Vibration of neck skin on larynx was detected by contact microphone (Lc). Sound radiating from nares was detected by small microphone (Nm). Sound radiating from nares and mouth was detected by standard microphone (NOm). The accelerometer system was used as standard, consisting of two pickups, one of which detected vibration of nasal skin (Na), and the other, neck skin on larynx (La). Nv/Lc, Nv/NOm, Nc/Lc, Nm/NOm were compared with Na/La. These values were normalized as the value for sustained /m/ as 1.0. Clinical use of these new methods was discussed.

2pSC7. Glottal characteristics of children—acoustic measures and comparison with adults. Helen M. Hanson, Kenneth N. Stevens (Sensimetrics Corp., 48 Grove St., Ste. 305, Somerville, MA 02144-2500, hanson@sens.com), and Ralph N. Ohde (Vanderbilt Univ., Nashville, TN 37212)

Previous studies of vowels produced by adult speakers showed substantial individual and gender differences in acoustic parameters related to glottal characteristics [H. M. Hanson, J. Acoust. Soc. Am. **101**, 466–481 (1997); H. M. Hanson and E. S. Chuang, *ibid.* **106**, 1064–1077 (1999)]. The current study extends the earlier work to children. In pilot work, data produced by seven children (five male, two female, ages 5-8) showing no signs of speech or voice disorders were analyzed. Acoustic measures related to first-formant bandwidth, open quotient, and spectral tilt were made on the vowel /ɛ/ extracted from the word ''red.'' These preliminary data show, for example, that the average measure of spectral tilt (dB difference between amplitudes of first harmonic and third-formant prominence) for the children is about 14 dB greater than for adult females and about 24 dB greater than for adult males. The results suggest that, like adults, children display substantial individual differences in vocal-fold configuration. Data for additional subjects and vowels will be analyzed, and we will make measures related to aspiration noise. Mean data for children and adults will be compared and interpretation in terms of acoustic models will be proposed. [Work supported by NIH Grant DC03744.]

2pSC8. Utterance–final voice quality variations: Their perceptual structure and acoustic correlates. Rebecca Herman (Speech Res. Lab., Indiana Univ., 1101 E. 10th St., Bloomington, IN 47405, reherman@indiana.edu)

Voice quality variations can be used to mark prosodic and discourse boundaries in running speech. In order to quantify acoustic events in the signal during the production of these voice qualities, it is necessary to first discover their perceptual structure. The perceptual structure can then inform the acoustic analysis and help to find perceptually salient acoustic correlates. In this study, a perceptual experiment was run in which the stimuli were single words spoken by multiple talkers, which had all been extracted from the ends of utterances. Pairs of stimuli were given similarity ratings by listeners, and the similarity ratings were used in a multidimensional scaling analysis. The dimensions found in the multidimensional scaling analysis were interpreted using acoustic analysis. The voice qualities of the female voices in this study were found to have a primary perceptual split which was interpretable acoustically as laryngealized versus nonlaryngealized, but the voice qualities found in the male voices did not group into categories, and the perceptual dimensions were not as clearly interpretable in acoustic terms.

2pSC9. Phonetic targets as the link between speech production and speech perception. Yi Xu (Northwestern Univ., 2299 N. Campus Dr., Evanston, IL 60208) and Emily Q. Wang (Rush Univ., Chicago, IL 60612)

We argue in this paper that the coherence between speech perception and production, which the motor theory maintains, lies in the common knowledge about speech production shared by the speaker and the listener. More specifically, we propose that abstract phonemic units are associated with phonetic targets that are articulatorily operable. These targets, how-

ever, are usually not fully achieved in speech production. Rather, they are often only approximated by different degrees under the limitation of various articulatory constraints. The process of speech production, therefore, is one in which various phonetic targets are being continuously approximated. As evidence, we will demonstrate with recent acoustic data how lexical tones in Mandarin are implemented through continuous approximation under specific articulatory constraints, how these constraints bring about extensive variations in the F0 contours, and how, despite the variations, tonal targets maintain their integrity in connected speech. We will then further argue that listeners, being also speakers themselves, should be able to discover the intended targets by noting the consistency the speaker manages to achieve despite the articulatory constraints. In other words, listeners hear not just what the speakers do, but more importantly, what they intend to do: We listen to hear what is being said.

2pSC10. Acoustic vowel space in 13 large Utah families. Kim Corbin-Lewis (Dept. of Communicative Disord. and Deaf Education, Utah State Univ., 1000 Old Main Hill, Logan, UT 84322-1000, KimC@coe.usu.edu), Julie M. Liss (Arizona State Univ., Tempe, AZ 870102), Hilary Coon (Univ. of Utah, Salt Lake City, UT 84113), and Steven D. Gray (Univ. of Utah Med. School, Salt Lake City, UT 84113)

Acoustic vowel space was examined in 131 speakers from 13 large Utah families ranging in age from 20-85 years with a median of 40 years. This population provides a unique opportunity to evaluate speech production characteristics in families with known genetic pedigrees and similar environmental influences. The purpose of the study was to determine whether acoustic manifestations of vowel production could be traceable to gender, age, or family membership. Results indicate that gender by itself explains 43.5% of the variance (highly significant, $p<0.0001$). Controlling for this gender effect, age is not a significant predictor of vowel space area ($p=0.29$). However, being in the same family is a significant predictor of vowel space area ($p=0.002$), independent of effects of gender and age. Being in the same family explains an additional 13.5% of the variance in area. [Work supported by NIH-NIDCD and the W. M. Keck Foundation.]

2pSC11. Individual differences in speech production: Voice-onset-time. J. Sean Allen and Joanne L. Miller (Dept. of Psych., 125 NI, Northeastern Univ., Boston, MA 02115, jallen@lynx.neu.edu)

As part of an ongoing study of talker-specific speech processing, we sought to document the existence of individual talker differences in voice-onset-time (VOT). Four female and four male talkers, all native speakers of American English, were asked to produce 30 repetitions each of 18 different words beginning with /p/, /t/, or /k/. Words were displayed visually on a computer monitor one at a time in a randomized order at a constant presentation rate, and talkers were asked to produce the words at a comfortable speaking rate. The VOT and duration of each spoken word were measured. The eight talkers showed substantial individual variation in speaking rate (as measured by word duration) and, as would be expected, slower speaking rates were associated with longer VOT values. In order to determine whether there were individual differences in VOT beyond those associated with speaking rate, we conducted analyses that controlled for speaking rate on a word-by-word basis. These analyses revealed cases of consistent individual differences in VOT. We are currently examining whether the structure of listeners' perceptual voicing categories reflects such individual talker differences. [Work supported by NIH/NIDCD.]

2pSC12. Perception of speaker characteristics with long and short samples. Blas G. Payri (LIMSI–CNRS, 91403 Orsay Cedex, France)

In this article, the perception of speaker's characteristics is compared using long and short samples. The sound material was an Italian read sentence (seicento-cinquantotto) and the last two syllables ''to'' from this sentence. Twenty healthy Italian native speakers (10 females, 10 males) read the sentence, with the same recording conditions. In the first experiment, the listeners performed a free categorization of the syllables and then of the sentences. Then, the listeners were asked to give qualifiers to describe the classes they had made. The aim was to obtain the main criteria for the description of long samples (sentences) and short samples (isolated syllables). For the sentences, the gender and prosody criteria were more frequently used by listeners, whereas the pitch and voice quality were preferred to classify the syllables. In the second experiment, listeners rated the same samples as previously, along predefined axes (age, gender, breath, liveliness, etc.). The ratings were compared with acoustic measures. The results show that the ratings for the whole sentence, and for the extracted syllables, can be significantly different, leading to the conclusion that, within a sentence, there are significant local variations of the perceived voice quality and speaker characteristics as gender and age.

2pSC13. Effects of listener experience on measures of voice quality. Jody Kreiman, Bruce R. Gerratt (Div. of Head and Neck Surgery, UCLA School of Medicine, 31-24 Rehab. Ctr., Los Angeles, CA 90095-1794, jkreiman@ucla.edu), and Melissa Epstein (UCLA, Los Angeles, CA 90095)

Previous studies have shown that clinicians and naive listeners use different perceptual strategies when judging the quality of pathological voices. Naive listeners uniformly rely almost exclusively on $F0$, while clinicians differ substantially from each other in the features they attend to. Consistent differences between groups in the overall level of rating have also emerged in paired comparison tasks, with naive listeners hearing greater differences between stimuli than clinicians do. We hypothesized that these differences in voice ratings are due to task-related factors, rather than to true differences in the ability of the two listener groups to hear differences in voice quality. Analysis-resynthesis approaches to quality measurement should eliminate such differences in measures of quality, because they require direct auditory matching of stimuli, without reference to internal standards that can be modified by auditory experience. To test this hypothesis, clinicians and naive listeners matched synthetic voice tokens to natural stimuli by adjusting synthesizer parameters of voice quality. No significant differences emerged between groups, suggesting that previously reported group differences are artifacts of traditional quality measurement tasks. Analysis-resynthesis approaches to measuring voice quality perception thus appear to control variability in judgments resulting from differences in internal standards. [Work supported by NIDCD Grant DC01797.]

2pSC14. The effects of speech rate and listening experience on the intelligibility of synthesized speech. Carol McCole (New York City Board of Education, Technology Solutions, District 75, 400 First Ave., New York, NY 10010 and Dept. of Speech, Commun. Sci., and Theatre, St. Johns Univ., Jamaica, NY 11439) and Fredericka Bell-Berti (St. Johns Univ., Jamaica, NY 11439)

Some persons with severe communication disorders are only able to communicate effectively using speech synthesizers, although even the most intelligible synthesis systems (e.g., DynaVox-DECTalk) have limitations. Listener success has been reported to vary with listener age, listener experience with synthesized speech, speech rate, and the presence of pauses between words. This series of experiments has examined the effects of listener age and experience, and speech rate and interword pauses on the perception of synthesized speech. Eighty IEEE-Harvard Sentences were presented to four groups of 12 listeners each [children (9–11 years), teens (14–16 years), young adults (20–25 years), and adults (38–45 years)], in two experiments. In the first experiment (40 sentences) speech rate was varied (105, 135, 165, and 195 wpm); in the second experiment (40 sentences), 7–12 days later, interword pauses occurred in half the sentences. Conditions were counterbalanced; no sentence was presented twice. Preliminary results indicate little effect of speech rate but a substantial effect of experience for the children: they improved substantially

through experiment 1; the improvement carried over into experiment 2. Parallel results will be presented for all age groups. [Work supported by St. Johns University and New York City Board of Education, Technology Solutions, District 75.]

2pSC15. The role of speaker familiarity in assessment of dysarthric speech intelligibility. Kuo-You Huang (Dept. of Childhood Education and Nursery, Chia-Nan Univ., Tainan, Taiwan, ROC)

It is well known that dysarthric speech is often more intelligible to familiar listeners. The underlying mechanisms for this speaker familiarity effect, however, are not fully understood. The present study was conducted to investigate effects of speaker familiarity and laboratory training on perception of dysarthric speech. In the first experiment, a Chinese-speaking woman with cerebral palsy produced lists of bisyllabic words, pseudowords, and isolated monosyllables. Familiar and naive listeners transcribed these items. Results showed that familiar listeners yielded substantially higher intelligibility scores, and that the effect was more pronounced for meaningful words than pseudowords and monosyllables. It was shown that familiarity effect resulted from the interaction between speech decoding and word knowledge. In the second experiment, different intensive training programs were carried out to discover the dynamics of the familiarity effect: (1) feedbacks about the target words were or were not given to listeners; and (2) a forced-choice paradigm was applied in the training task and confusability among syllable pairs was manipulated, so that top-down and bottom-up influences could be differently observed during the training course. Results indicated that laboratory training was effective and it can, in fact, mimic the development of the familiarity effect in the naturalistic environment.

2pSC16. Intonation and emotion. Marc D. Pell (School of Commun. Sci. and Disord., McGill Univ., 1266 Pine Ave. W., Montreal, QC H3G 1A8, Canada, mpell@po-box.mcgill.ca)

Preliminary data were gathered on how simulated emotional characteristics of the voice influence the acoustic form of English utterances containing specific combinations of intonational and stress features [extending Eady and Cooper, J. Acoust. Soc. Am. **80**, 402–415 (1986)]. Utterances varying in intrasentential focus position (initial, final, none) were elicited as both statements and questions in each of four emotional ''modes'' (neutral, sad, happy, angry) employing a structured elicitation procedure. Parameters of duration and fundamental frequency ($f0$) were then determined for the productions elicited from eight elderly speakers to specify important acoustic dimensions associated with specific combinations of stress, ''modality,'' and emotional features of the stimuli. Results of the acoustic analyses largely reaffirmed past accounts of how contrastive focus is encoded in (affectively neutral) statements and questions for English, and cohered well with the acoustic literature on how basic emotions are expressed vocally for three key acoustic parameters (mean $f0$, $f0$ range, speech rate). The impact of emotion on linguistic attributes of prosodic structure was most evident in the speakers' modulation of $f0$, which was notably constrained in prosodic conditions where speakers were required to signal ''marked'' emotional and nonemotional intentions conjointly within the intonation contour. [Work funded by FCAR.]

2pSC17. Eliciting speech styles in the laboratory: Assessment of a new experimental method. James D. Harnsberger, David B. Pisoni (Speech Res. Lab., Dept. of Psych., Indiana Univ., Bloomington, IN 47405, jharnsbe@indiana.edu), and Richard Wright (Univ. of Washington, Seattle, WA 98195)

This study concerned the development of a method of eliciting controlled stimulus materials, namely sentences, varying in three speech styles: reduced, citation, and hyperarticulated. Such a method would prove useful in studying the acoustics and perception of speech styles. The reduced style was elicited by having subjects read sentences while carrying out a distractor task that involved recalling a digit sequence from short-term memory. The length of the sequence was calibrated to the individual subject via a digit span task administered prior to elicitation. It was hoped that the distractor task, serving as a cognitive load, would result in a less monitored, more naturalistic style. The citation style was elicited by having subjects read sentences from a list. The hyperarticulated style was elicited in a similar manner, except that subjects were prompted on some trials to reread the sentence again more carefully. Twelve subjects were recorded in this experiment, and the resulting sentences were acoustically analyzed and perceptually tested in a pairwise comparison task. The results showed that half of the talkers produced reduced sentences that differed acoustically and perceptually from citation sentences, while all 12 subjects produced hyperarticulated sentences that were differentiable from all others.

2pSC18. Direct classification of Thai monophthongs on two-dimensional acoustic–phonetic feature spaces in linear, mel, bark, and bark-difference frequency scales. Visarut Ahkuputra, Somchai Jitapunkul, Ekkarit Maneenoi (Digital Signal Processing Res. Lab., Dept. of Elec. Eng., Faculty of Eng., Chulalongkorn Univ.), and Sudaporn Luksaneeyanawin (Chulalongkorn Univ.)

A direct classification of nine Thai monophthongs was conducted on two-dimensional acoustic–phonetic feature spaces. Formant frequencies of a monophthong have been extracted from a stable vowel portion of a syllable nucleus. The $F1$ and bark-difference $F1$-$F0$ represents tongue height. The $F2$ and bark-difference $F3$-$F2$ represents tongue advancement. The Bayesian classifier utilizes statistical parameters, mean and standard deviation, of each monophthongs in classification. Two-dimensional feature vectors comprise two acoustic–phonetic features, $F2$ and $F1$ in linear, mel, bark, and also, bark-difference $F3$-$F2$ and $F1$-$F0$. The classification results on the $F2$ and $F1$ space are 86.3325% in linear, 84.6187% in mel, 84.2331% in bark, 79.3916% in bark-difference $F2$-$F1$ and $F1$-$F0$, and also 79.9914% in bark-difference $F3$-$F2$ and $F1$-$F0$. Considering confusion matrices, the high vowels (/i:, v:, u:/), front-middle vowel (/e:/), and center-middle vowel (/q:/) have higher results than other groups which resulted from smaller overlapping and more separability between each group compared to low- and back-vowel groups. The back vowel (/u:, o:, @:/) has more misclassification between each monophthong than in other groups. Both feature spaces in bark-difference frequency scales exhibit lower-classification results although they illustrate better vowel distribution than other scales.

Session 2pUW

Underwater Acoustics and Acoustical Oceanography: High-Frequency Sediment Acoustics and Associated Sediment Properties and Processes II

Kevin L. Williams, Chair
Applied Physical Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

Chair's Introduction—12:55

Contributed Papers

1:00

2pUW1. The scattering of low-grazing angle pulse beams at interfaces between fluid and porous media. Ralph A. Stephen (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Many environments in bottom interacting ocean acoustics consist of mud and sands which can be modeled as fluid saturated, porous solids with low shear moduli. To study the physical mechanisms responsible for forward- and backscattering in these media it is useful to have a forward modeling technique which applies to rough and heterogeneous porous bottoms. A numerical scattering chamber using the time-domain finite-difference method applied to the range-dependent Biot equations can be used to study the scattering of low-grazing angle pulse beams from fluid-saturated porous media. For example, for a flat seafloor over a homogeneous porous half-space, converted shear waves and converted compressional ''slow'' waves are shown to be excited in the subbottom even when the grazing angle is below the critical angle for compressional ''fast'' waves. Any scattering element near the seafloor will act as a secondary point source by Huygen's principle and when excited by an acoustic wave will have the potential to generate a family of interface waves and body waves of ''fast'' compressional, ''slow'' compressional and shear type.

1:15

2pUW2. At-sea measurements of sound penetration into sediments using a buried vertical synthetic array. Harry J. Simpson, Brian H. Houston (Naval Res. Lab., Code 7136, 4555 Overlook Ave., Washington, DC 20375-5320), Alian R. Berdoz, Philip A. Frank, Steve W. Liskey (SFA, Inc., Largo, MD 20774-5322), J. S. Stanic (Stennis Space Center, MS 39529-5004), and Carl K. Frederickson (Univ. of Central Arkansas, Conway, AR 72035-0001)

At-sea measurements were conducted to understand the propagation of sound into sediments. These investigations used a new, buried vertical synthetic array system developed and built at the Naval Research Laboratory. The design of this at-sea system is based on a system developed for laboratory experiments that utilized a two-dimensional synthetic array technique [J. Acoust. Soc. Am. **107** (2000)]. For the measurements reported here, the hydrophone was water-jetted into the sediment to a depth of 2 m. Once buried, the hydrophone was mounted to a vertical robotics stage that pulled the hydrophone up in 1-cm increments. A spherical source, positioned 50 cm above the interface, was used to ensonify the sediment. The ensonification angle was varied by changing the horizontal distance of the source from the insertion point. In this manner, several measurements were made with ensonification angles above and below the critical angle. In addition, the enhancement of sound penetration due to interface roughness was investigated. This new system, along with the results of the measurements, will be discussed. [Work supported by ONR.]

1:30

2pUW3. Spatial and temporal variation of large- and small-scale seafloor roughness. Kevin B. Briggs (Seafloor Sci. Branch, Naval Res. Lab., Stennis Space Center, MS 39529-5004)

Seafloor microtopography, measured at subcentimeter resolution over scales up to a meter or more by underwater stereo photogrammetry, was periodically characterized during the SAX99 high-frequency acoustic backscattering experiment. Analog 35-mm stereo photographs of seafloor roughness were digitized at 2- to 5-mm intervals from film transparencies with a stereo comparator capable of 1-mm-scale lateral precision and submillimeter-scale vertical accuracy. Relative seafloor height profiles were used to estimate roughness power spectra, which are parametrized as slopes and intercepts of regressions fit to the roughness power spectra. The composite roughness model developed by Applied Physics Laboratory–University of Washington requires these two spectral parameters to determine the contribution of the seafloor roughness to the overall backscattering strength. Data from two locations over 31 days in the experimental area off Ft. Walton Beach, FL in 19-m water depth are presented in order to demonstrate the relationship between seafloor roughness morphology at subcentimeter to meter scales and model parameters predicting high-frequency acoustic scattering. Model predictions are generated from the measured geoacoustic properties of the rippled, medium sand seafloor and the range of roughness parameters that varied temporally and spatially during the experiment. [Work supported by ONR.]

1:45

2pUW4. High-resolution measurements of volume heterogeneity in sandy sediments. Robert A. Wheatcroft, Cara E. Fritz, and Louise Hunt (College of Oceanic and Atmospheric Sci., Oregon State Univ., Corvallis, OR 97331)

As part of SAX 99, a multi-investigator study focused on the penetration and scattering of high-frequency (10–300 kHz) sound in marine sediments, two independent measurements of sediment volume heterogeneity (SVH) were made. The first tool used to measure SVH was an *in situ* resistivity profiler (IRP) that logs sediment resistivity at 0.25-mm vertical intervals to depths of 15 cm. Robust empirical calibrations are used to calculate sediment porosity. The second method used to measure SVH was a digital x-ray system that comprises a constant-potential x-ray source and an amorphous silicon x-ray detector. The IRP data indicate little between-site variability and bottom-type (i.e., ripple crest, ripple trough, mound) specificity. Profiles are characterized by a zone from 0- to 3-mm subbottom of exponentially decreasing porosity, underlain by a zone of roughly constant porosity characterized by 5%–10% fluctuations. The digital x radiographs indicate three sources of SVH: (1) biogenic structures, (2) mud inclusions, and (3) shell fragments. Of these, the shell fragments are likely to be the most important sound scatterers. Estimates of shell fragment concentrations, distributions, and orientations will be discussed. [Work supported by ONR.]

2:00

2pUW5. APL-UW environmental measurements during SAX99: Sediment conductivity and tomography. Dajun Tang, Kevin L. Williams, Darrell R. Jackson, and Eric I. Thorsos (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

Two new instruments were developed and deployed during SAX99 to measure surficial sediment variability at centimeter scales. Such data serve as input to acoustic models predicting sound scattering in the frequency range of 10–50 kHz. One instrument, IMP (In situ Measurement of Porosity) measures sediment conductivity at 1-cm resolution in the horizontal dimensions and at 3-mm resolution in the depth dimension. From this instrument the following information is derived: (1) 3-D porosity or density variation in the top 12 cm of sediments, and (2) 2-D bottom roughness and associated spectra. The second instrument, the Acoustic Imager (AI), is a 3-D sediment tomographic tool with 1-cm resolution operating at 170 kHz. Information derived from the AI includes (1) 3-D sediment sound speed variability, (2) 3-D variability of sediment attenuation coefficients, (3) the presence and distribution of discrete scatterers such as shell pieces, and (4) the temporal variability of the above parameters over 3 days. These results and their implications to the acoustic measurements taken during the SAX99 experiment will be discussed. [Work supported by ONR.]

2:15

2pUW6. Quantification of Darcy's coefficient of permeability in siliciclastic marine sands using image analysis and effective medium theory modeling. Allen H. Reed and Dawn Lavoie (Naval Res. Lab., Bldg. 1005, Stennis Space Center, MS 39529)

Darcy's coefficient of permeability (Dcp) and porosity strongly influence the behavior of sound in unconsolidated sediments at high frequencies, therefore it is important to assess physical property values in sediments undisturbed by coring and subsequent handling. The objective of this study was to determine Dcp and porosity of siliciclastic marine sands from Ft. Walton Beach, FL that were embedded *in situ* using a polyester casting resin. For comparison, direct measurements of Dcp were made with an *in situ* falling head permeameter and porosity was determined by water weight loss. Modeling Dcp entailed quantifying pore dimensions and pore connectivity in the resin-embedded medium from two-dimensional scanning electron microscopy images using standard image analysis techniques. The value of Dcp is estimated from average, or ''effective,'' pore characteristics using an effective medium theory technique. Dcp of the samples in both vertical and horizontal directions averaged $1.9 \times 10\text{-}3 \pm 1.3 \times 10\text{-}3$ and porosity averaged 0.44 ± 0.04. The modeled and measured Dcp values are comparable, whereas modeled porosity was slightly higher than measured porosity. From these data we conclude that estimating Dcp and porosity from *in situ* embedded sands results in values of physical properties close to those in the undisturbed state. [Work funded by ONR and NRL.]

2:30–2:45 Break

2:45

2pUW7. Estimation of seafloor roughness using ROV-mounted lasers. Brett K. McCurley, James N. Piper, and Nicholas P. Chotiros (Appl. Res. Labs., Univ. of Texas, Austin, P.O. Box 8029, Austin, TX 78713-8029, chotiros@arlut.utexas.edu)

The method employs lasers mounted on the remotely operated vehicle (ROV). Using a cylindrical lens to spread a laser beam into a fan and viewing the intersection of the laser fan with the water-sediment interface, it is possible to obtain a profile of the sediment. As the ROV advances, a bathymetric profile of the sediment may be constructed using successive profile snap shots. Data were collected at the Sediment Acoustics Experiment (SAX99) and processed to produce sediment surface profiles. [Work supported by the Office of Naval Research, Ocean Acoustics.]

3:00

2pUW8. Mechanisms for biological modification of seabed: Direct acoustic observations. D. V. Holliday, C. F. GreenlawIII, and D. E. McGehee (BAE Systems, 4669 Murphy Canyon Rd., Ste. 102, San Diego, CA 92123-4333)

Surface roughness and volume heterogeneity have both been implicated in the anomalous penetration of high-frequency sound into the ocean bottom. These physical characteristics of the seabed are dynamic properties of the bottom environment, with changes driven by both biotic and abiotic forces. Species that use the bottom for shelter from visual predators can change surface roughness at the time of emergence and re-entry. Acoustic observations of the movement of marine endo- and epibenthic organisms between the seabed and the water column will be used to illustrate the potential for these behaviors to modify the bottom. This behavior can also be the source of some of the dynamic volume heterogeneity that characterizes the seabed. The maintenance of burrows by animals is a known source of changes in both surface roughness and volume heterogeneity. At one study site fish appeared to anticipate the behavior of the plankton, foraging on lower trophic levels near, in, and on the seabed. This behavior can also modify surface roughness at the bottom-water column interface. [Research supported by ONR.]

3:15

2pUW9. Assessing 3D volume inhomogeneity using a noninvasive diver-deployed resistivity array. Peter D. Jackson, David A. Gunn (Br. Geological Survey, Keyworth, Nottingham NG12 5GG, UK), and Kevin B. Briggs (Naval Res. Lab., Stennis Space Center, MS 39529)

A noninvasive electrical resistivity array imaging system (RAIS) has been developed to measure sediment volume inhomogeneities in three dimensions. Novel measurements of electrical fields, due to controlled galvanic current-flow, have been inverted into electrical resistivities, in three dimensions beneath the seafloor. The limitations of conventional 3D resistivity inversions are explored, and additional *a priori* constraints are shown to significantly improve performance in the seafloor environment. From these 3D resistivity distributions, correlation lengths in x, y, and z dimensions are calculated to support the modeling of acoustic scattering from the seafloor. Examples obtained during November 1999 in the Gulf of Mexico for the ONR High-Frequency Sound Interaction in Ocean Sediments (SAX99) experiment are described. New technology has reduced the acquisition time for a 3D dataset of 3840 measurements from approximately 2 h to 90 s, enabling a suite of locations to be investigated during a single dive. Converting to formation factors and combining with porosity/density data enabled to tortuosity of the electrical current-paths to become a pore morphology descriptor. Formation factors were found in the range 2 to 5, being consistent with substantial bioturbation in sandy sediments, with heterogeneities identified up to 100 mm.

3:30

2pUW10. Imaging the permeability structure within near-surface sediments by acoustic crosswell tomography. Tokuo Yamamoto (Appl. Marine Phys. Div., RSMAS, Univ. of Miami, Miami, FL 33149)

Permeability structure within near-surface sediments is imaged by analytically inverting the measured fields of acoustic wave velocity and attenuation. The acoustic velocity and attenuation fields at 1000 Hz were previously measured by a pseudorandom binary sequence (PRBS) based crosswell acoustic tomography device [T. Yamamoto, J. Acoust. Soc. Am. **98**, 2235–2248 (1995)]. A quadratic equation of permeability derived from the Biot theory is used for the permeability imaging. The porosity and shear modulus images are also extracted from the velocity and attenuation fields using the newly derived quadratic dispersion equation and an empirical shear modulus porosity depth of burial relation. The acoustically measured images of permeability, porosity, and shear modulus are favorably compared with the engineering tests performed on the cores and boreholes.

3:45

2pUW11. Ultrasonic studies on the effect of water content on compressional wave velocity in beach sand. Jacques R. Chamuel (Sonoquest Advanced Ultrason. Research, P.O. Box 81153, Wellesley Hills, MA 02481)

Theoretical results by Bachrach and Nur [Geophysics **63**(4), 1225–1233 (1998)] indicate that the compressional wave velocity in wet sand decreases from 1700 to 250 m/s as the moisture content is decreased by 1% from saturation based on the Biot-Gassmann theory low-frequency limit. New quantitative experimental results from beach sand (280 μm average grain size) are presented. Experimental results from wet sand ''half-space'' are compared with small wet sand samples. The sand moisture content was continuously measured from the weight of a mechanically balanced fixture holding ultrasonic transducers and the wet sand sample. Preliminary quantitative experimental results from the mechanically balanced wet sand show that the compressional wave velocity remained near 1650 m/s in the 100–500 kHz frequency range as the moisture content was decreased by more than 35% from saturation (typical drained sand is only about 5% from saturation). Beach sand, being angular with flat grain areas, may need to be modeled as platelets with large contact areas rather than spherical grains with point contact.

4:00

2pUW12. Effects of echo variability in geoacoustic sediment characterization. Daniel D. Sternlicht (ORINCON Corp., 9363 Towne Ctr. Dr., San Diego, CA 92121, dsternlicht@orincon.com) and Christian P. de Moustier (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0205)

The variability in shape and amplitude of bottom echoes affects the resolution of geoacoustic parameter estimation techniques that match measured bottom-echo envelopes with envelope models derived from incoherent acoustic backscatter theory at the sediment–water interface and in the sediment volume. Following earlier echo envelope matching work carried out with data collected at 10–100 kHz in San Diego Bay [Sternlicht and de Moustier, J. Acoust. Soc. Am. **105**, 1206 (1999)], coarse- and fine-grain substrates are identified by estimates of their mean grain size (M_ϕ) and of the strength (w_2) of their power-law roughness spectrum describing the spatial statistics of the interface. Statistics of the (M_ϕ, w_2) estimates are combined with those from Monte Carlo simulations, based on the data covariance matrices, to demonstrate that in sandy substrates, echo variability has little effect on estimates of w_2, and a pronounced effect on estimates of M_ϕ. The reverse pattern is observed for fine-grain substrates. However, the parameter estimation technique used introduces a degree of correlation between M_ϕ and w_2, which is especially large for sand substrates. These observations are consistent with nature, where relief spectra of coarser sediments contain more energy than those of fine-grain substrates. [Work supported by ONR.]

4:15

2pUW13. Application of the Gelfand–Levitan method to geoacoustic inversion in shallow water. Kyle M. Becker (MIT/WHOI Joint Prog. in Oceanogr. and Oceanogr. Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543) and George V. Frisk (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

The Gelfand–Levitan inverse technique is an exact inversion method originating in quantum scattering theory [I. M. Gelfand and B. M. Levitan, Am. Math. Soc. Transl. **1**, 253–304 (1955)]. At a single frequency, the required input data are the Fourier transform of the plane wave reflection coefficient as a function of incident vertical wave number. The method was originally employed in underwater acoustics for a deep-water environment where a simple relationship exists between the reflection coefficient and the depth-dependent Green's function, which is obtained by Hankel transforming measurements of the point source pressure field [A. A. Merab, Sc.D. thesis, MIT/WHOI Joint Program, Woods Hole, MA (January 1987)]. In shallow water, the Green's function is characterized by poles corresponding to the eigenvalues of the perfectly trapped modes in the waveguide. These poles complicate the relationship between the Green's function and the reflection coefficient, and therefore the application of the Gelfand–Levitan technique. However, through application of the Darboux transform to the governing wave equation and the reflection coefficient, the Gelfand–Levitan method can be employed for shallow-water geoacoustic inversion [J. R. McLaughlin and S. Wang, in *Mathematical and Numerical Aspects of Wave Propagation*, edited by J. A. DeSanto (SIAM, Philadelphia, 1998), pp. 232–236]. In this talk, the Darboux transform will be illustrated and examples of geoacoustic inversion for synthetic shallow-water acoustic data will be presented. [Work supported by ONR.]

4:30

2pUW14. Geoacoustic parameters by inversion of bottom reflectivity data from the Santa Barbara channel experiment. Ross Chapman and Danya Hudson (Ocean Acoust. Group, School of Earth and Ocean Sci., Univ. of Victoria, P.O. Box 3055, Victoria, BC V8W 3P6, Canada, chapman@uvic.ca)

A method is described to estimate the range dependence of geoacoustic parameters of the seafloor from bottom reflectivity data. The data were obtained using light bulbs as sound sources in the Santa Barbara channel experiment (SBCX). Light bulbs were deployed around the SBCX FFP array that consisted of five 30-element vertical line arrays. Signals from the direct path, the surface reflected path, and the four components of the first bottom reflection were clearly resolved at each array. The inversion integrates the results from several source–array combinations to generate estimates of the seafloor geoacoustic parameters in the area around the FFP array. The reflection coefficients were determined versus angle from the direct path and bottom-reflected path signals, and inverted using a simple fluid model to estimate the seafloor sound speed and density. The estimates are consistent with the known ground truth information for the SBCX site. [Work supported by ONR.]

4:45

2pUW15. Sediment tomography in the East China Sea. Gopu R. Potty and James H. Miller (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett, RI 02882)

Sediment tomography experiments using broadband shots are planned as part of the Asian Seas International Acoustic Experiment (ASEAEX) in the East China Sea (ECS) in the year 2001. In order to gather adequate sediment information, gravity cores were taken from the experimental location during the spring of 2000. Using the gravity core data and available historic data, a sediment geoacoustic model for the region is presented. Synthetic experiments are performed with this geoacoustic model using broadband data. Various configurations of the explosive sources are tested to arrive at a suitable experimental geometry for the tomography experiment. This inversion is based on group speed dispersion of the broadband acoustic signal and is capable of estimating range-dependent sediment properties. [Work sponsored by ONR.]

Session 3aAA

Architectural Acoustics: Integration of Synthesis Techniques and ''Acoustical'' Music

Richard H. Campbell, Cochair
Bang-Campbell Associates, Box 47, Woods Hole, Massachusetts 02543

K. Anthony Hoover, Cochair
Cavanaugh Tocci Associates, Inc., 327F Boston Post Road, Sudbury, Massachusetts 01776

Chair's Introduction—8:25

Invited Papers

8:30

3aAA1. The real time use of signal processing in brass performance. Thomas J. Plsek (Berklee College of Music, 1140 Boylston St., Boston, MA 02215)

The real time use of signal processing for brass instrument performance (specifically the trombone) provides an opportunity, as well as challenges, to create a type of hyper-instrument quite unlike most conceptions of the term. These include the expansion of the available sound palette, the problem of real time parameter control, microphone techniques, and repertoire considerations. The various types of processing used includes reverb, pitch change, delays—short time and longer looping types, distortions, and combinations of the above. One of the notions that must be accepted is that, unlike instruments whose sound is produced electronically, the processed signal, unless significantly long-delay-based processing is used, is always layered onto an acoustic output. By creatively managing the acoustic instrument and the electronic equipment, a vast array of musical resources become available to the performer enabling him/her to enhance existing performance environments, as well as find and develop new ones.

8:50

3aAA2. New control technologies for virtual orchestra playback. David B. Smith (New York City Tech. College, 186 Jay St., Brooklyn, NY 11201), Kojiro Umezaki (Healthy Boys, L.L.C., New York, NY 10010), Frederick W. Bianchi, and Richard H. Campbell (Worcester Polytechnic Inst., Worcester, MA 01609, rhcamp@ma.ultranet.com)

A new software application running on the Linux operating system has been developed exclusively to support live performance by a virtual orchestra (VO). The core engine, which allows unlimited expansion, has a rich feature set including temporal flexibility, arbitrary performance navigation, printed score compliance, multiple input/output devices, and editing capability. The base configuration supports 64 MIDI channels and utilizes metaevent extensions to the MIDI 1.0 specification. MIDI files from multiple platforms can be imported. Considerable attention has been given to human factors in VO real-time music production with meaningful display screens and specialized keyboard designs resulting in a rapid and intuitive user learning curve. [Work supported by Realtime Music Solutions, L.L.C., New York.]

9:10

3aAA3. Synthesized space in pop music production. Alexander U. Case (Cavanaugh Tocci Assoc., Inc., 327F Boston Post Rd., Sudbury, MA 01776, acase@cavtocci.com)

The sonic vocabulary of the pop music recording engineer is not constrained by the architectural acoustics of any space. Equipped with racks and RAM full of signal processing tools, the engineer tries to create musical timbres, textures, ambience, and emotions to support whatever feeling the music inspires. Perhaps nowhere is this departure from the physical to the contrived more apparent than in reverb. Acoustic analogies like the chamber, mechanical simulations like the plate, and the wholly invented spaces of the digital reverb device are part of a family of effects which—inspired by the great halls and opera houses of the world—have an identity and freedom all their own. Reverb time, early decay time, bass ratio, initial time delay gap, and so on become independent variables freely manipulated by the engineer to create sonic ''spaces'' that may not be physically possible outside of the studio. This paper surveys contemporary pop music production trends in reverb and analyzes them through the lens of the architectural acoustician.

9:30

3aAA4. Peformances of Ballet Mécanique. Paul D. Lehrman (Tufts Univ., Medford, MA)

George Antheil's 1924 Ballet Mécanique, written for two pianos, three xylophones, four bass drums, tam-tam, electric bells, three airplane propellors, and between 4 and 16 synchronized player pianos, is one of the great ''lost'' pieces of the 20th-century instrumental repertoire. This paper describes the efforts that were made between 1996 and 2000 to revive this remarkable piece, using modern computers, MIDI, digital samplers, and MIDI-compatible player pianos. It talks about how the project same about; the musical and technical decisions that needed to be made; the problems of transcription, synchronization, and rehearsal; and the practical, artistic, and logistical issues of presenting the piece in three different concert spaces: Durgin Hall at the University of Massachusetts Lowell, Carnegie Hall in New York, and Louise Davies Symphony Hall in San Francisco.

9:50–10:05 Break

10:05

3aAA5. Quantitative practice for students at the Digital Conducting Laboratory. Teresa Marrin Nakra (Immersion Music and Worcester Polytech. Inst., 381 Farrwood Dr., Ward Hill, MA 01835, teresa@immersionmusic.org) and Gary Hill (Arizona State Univ.)

Students of conducting and conducting pedagogy at Arizona State University's School of Music now have the opportunity to enhance their skills as performers and teachers with the assistance of a unique technology housed in the Digital Conducting Laboratory. Utilizing a Digital Conducting Feedback System designed by Dr. Teresa Marrin Nakra of Immersion Music, Inc., the laboratory has the capability to simulate many of the behaviors of a live orchestra rehearsal setting. The system's unobtrusive sensor interface and interactive program recreates several fundamental ensemble-conductor interactions, particularly by reacting to the tempo, articulation, and dynamic line generated by the conductor. In addition to immediate aural feedback, the system allows conductors to review their performances via sound files, video playback, and analysis of muscle-tension profiles. The musical materials consist of a set of etudes that systematically take a conductor through a review of basic conducting gestures. Compiled and arranged by the laboratory's founder and director, ASU Professor Gary W. Hill, the etudes are principally derived from familiar classical music, freeing the user to focus primarily on his or her conducting and the concomitant response.

10:25

3aAA6. Interactions among acoustics, digital signal processing, and movement. Seth Cluett (64 Beacon St., Boston, MA 02108, sethcluett@yahoo.com)

A team of architects, electronic artists, and acousticians designed a system for manipulating acoustic, haptic, and spatial information based on the motion of dancers through space. The dance structure was designed to inspire physical exploration and interaction, to sense and process resultant sounds based on this dance, and to present an immersive sound field to the dancers and observers which complemented the visual impact of the movement. A microphone system captured the sounds of the dancers' contact with the system. An electronic sensor system tracked the location and action of the dancers. A set of digital signal processing algorithms analyzed, modified, and spatialized in real time the information collected from the environment, integrating sensory cues arriving simultaneously and rapidly via multiple modalities. A multichannel speaker system re-presented the natural and modified sounds to the acoustic environment, which in turn affected the dancers' movements. The nature of movement through physical space, the localization of movement in both vision and audition, and the abstraction, support, and deconstruction of both modalities were studied. The results were used to design a performance by a dance ensemble that explored the synergy between movement, multimodal sensory information, and sound.

10:45

3aAA7. Alternative voices for electronic sound. Dan Trueman (Computer Music Ctr., Columbia Univ., New York, NY 10027), Curtis Bahn (iEAR Studios, Rensselaer Polytechnic Inst., Troy, NY 12180), and Perry Cook (Princeton Univ., Princeton, NJ 08544)

Through the design and construction of unique sound diffusion structures, the nature of electronic sound can be reinvented. When allied with new sensor technologies, these structures offer alternative modes of interaction with techniques of sonic computation. This paper describes several recent applications of Geosonic Speakers (multichannel, outward-radiating geodesic speaker arrays) and Sensor-Speaker-Arrays (SenSAs: combinations of various sensor devices with outward-radiating multichannel speaker arrays). Geosonic Speakers, building on previous studies of the directivity of acoustic instruments (the NBody Project), attempt to reproduce some of the diffusion characteristics of conventional acoustic instruments; they engage the reverberant qualities of performance spaces, and allow electronic and acoustic instruments to blend more readily. In collaboration with the U.S. Enclosure Company, over a dozen Geosonic Speakers were produced, varying in size from 8 to 14 in. Their use in the performance and recording of two major new works by composer Steven Mackey is detailed. The electronic improvisation ensemble ''Interface'' has integrated a family of these Geosonic Speakers into their standard setup, completely replacing their previous P.A. diffusion model; this has encouraged them to substantially reinvent their approach to the performance of live interactive computer music.

Contributed Paper

11:05

3aAA8. Spectro-temporal control parameters for synthetic orchestra applications. James W. Beauchamp (School of Music and Dept. of Elec. and Computer Eng., Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801, j-beauch@uiuc.edu)

Several spectral properties can be exploited for control of individual voices in a synthetic orchestra. Sounds can be synthesized from their spectral representations using time-varying additive synthesis, inverse Fourier transform, or, more efficiently, using multiple wavetable synthesis. Primary parameters to be controlled are amplitude and spectral centroid versus time envelopes, spectral envelope, spectral irregularity, vibrato, noise content, and inharmonicity. These parameters can be further resolved into subparameters. For example, amplitude versus time curves can be partitioned into attack, steady-state, and decay segments which can be manipulated separately. For sustained sounds, it is also important to consider how these parameters vary with fundamental frequency and intensity. In the case of percussion sounds, the density of partials as well as their spectral evolution are important. In all cases, detailed management of parameters is necessary to maintain timbral clarity, uniqueness, and naturalness. Examples of synthetic orchestral instruments will be played.

Session 3aABa

Animal Bioacoustics: Use of Acoustics for Wild Animal Surveys

David K. Mellinger, Chair
Cooperative Institute for Marine Resources Studies, Oregon State University, 2030 SE OSU Drive, Newport, Oregon 97365

Chair's Introduction—7:30

Invited Papers

7:35

3aABa1. Acoustic detection distances of sperm whales in the Gulf of Mexico. David K. Mellinger (Cooperative Inst. for Marine Resources Studies, Oregon State Univ., 2030 SE OSU Dr., Newport, OR 97365), Aaron M. Thode (MIT, Cambridge, MA 02139), Anthony Martinez, Keith Mullin (Southeast Fisheries Sci. Ctr., Miami, FL 33149), and Sarah Stienessen (Texas A&M Univ., Galveston, TX 77551)

During a cruise in June–July 2000 in the north-central Gulf of Mexico, sperm whales were acoustically detected, located, and tracked using a small-aperture hydrophone array towed at a depth of 20–100 m. Water depth was 800–2000 m. Detection ranges to groups of sperm whales were estimated on several occasions by determining the location of a calling whale or group of whales, then moving away in a straight line until the group was undetectable. The range at which sperm whales could no longer be detected, either aurally with headphones or visually in a spectrogram, was found to be 4–6 km, significantly less than the 11 km that had been estimated previously in the Gulf [Norris *et al.*, Minerals Management Service Report 1996-0027]. To investigate this result, acoustic propagation models were run. Results from the models are shown and are used to explain the short detection distances present. In addition, propagation models are used to evaluate the effects of water depth, sensor depth, and bottom composition on detection range, which will be most useful for future attempts to predict detection ranges of sperm whale sounds. [Work supported by ONR, MMS, and NMFS.]

7:55

3aABa2. Using one or two hydrophones for marine animal surveys. Douglas H. Cato (Defence Sci. and Technol. Organisation, P.O. Box 44, Pyrmont, NSW 2009, Australia) and Robert D. McCauley (Curtin Univ. of Technol., Bentley, Western Australia)

Acoustic surveying of marine animals requires the ability to localize the sources or at least determine their distances, if the density of the animals is to be determined. Precisely placed multiple-element arrays will provide this information, but there are logistically simple techniques using one or two hydrophones that are effective under certain conditions. If the propagation conditions are known, the differences in the times of arrival and the received levels on two hydrophones provide the distance of the source and the source level, for sources significantly closer to one hydrophone than to the other. The positions of the hydrophones do not need to be known, but if they are, the sources can be localized with left–right ambiguity. An example is the use of sonobuoys deployed from a ship in transit. One hydrophone is sufficient if the direct and surface reflected paths can be separated, since the surface image can be used in place of the second hydrophone. Some examples of this applied to movement of fish sources will be presented. Complex propagation causes difficulties but these may also be exploited to improve the estimates.

8:15

3aABa3. Multi-modal surveys of fin whales in the Sea of Cortez, Mexico. Christopher W. Clark, Don A. Croll, Alejandro Acevedo, and Jorge Urban-Ramirez (Cornell Lab. of Ornithology, Bioacoustics Res. Prog., 159 Sapsucker Woods Rd., Ithaca, NY 14850, cwc2@cornell.edu)

A population of fin whales (*Balaenoptera physalus*), resident to the Gulf of California, Mexico, was studied over two seasons using an integrated approach. Systematic vessel-based visual survey and photo-ID efforts were conducted every 5–7 days to independently estimate the number and distribution of whales within a 10×30 mi^2 study area. Some whales were tagged with time-depth-recorders. Sets of 5–6 distributed autonomous seafloor acoustic recorders, operating continuously during each season's research period, were used to detect, locate, and track vocalizing whales. A 16-element towed array tracked individual vocal whales in real-time concurrently with visual observations, allowing biopsy samples of known vocal animals. Active acoustics was used to collect data on the density and distribution of krill so as to place measured variation in whale numbers, distribution, and behavior within an ecological context. The primary whale activity was feeding. Whale feeding patterns and survey distribution followed prey distribution. Vocal whale distribution followed dial feeding patterns and prey distribution. All biopsied vocal animals were males. Numbers of whales estimated by vessel survey, photo-ID, and passive acoustics were correlated. Results suggest that under certain conditions, vocal activity is a reliable measure of distribution and relative abundance. [Work supported by ONR.]

8:35

3aABa4. Locating and enumerating endangered humpback whales in the eastern Caribbean with directional (DIFAR) sonobuoys. Steven Swartz, Phillip Clapham, Tim Cole, Jay Barlow (U.S. Dept. of Commerce, NOAA, Natl. Marine Fisheries Service, 75 Virginia Beach Dr., Miami, FL 33149, steven.swartz@noaa.gov), Mark McDonald (Whale Acoust., Laramie, WY), Erin Oleson, and John Hildebrand (Scripps Inst. of Oceanogr., San Diego, CA)

Visual surveys for whales are limited by visibility conditions at sea and whale diving behavior. Male humpback whales (*Megaptera novaeangliae*) sing complex ''songs'' during their winter breeding season, and low-frequency song components lend themselves to acoustic detection. A visual and acoustic survey was conducted during the peak breeding season to evaluate passive acoustic methods to assess endangered humpback whales in the Eastern Caribbean, a region where they were exploited to depletion. Directional (DIFAR) sonobuoys were used to detect singing whales. Bearing angles from the sonobuoys to singing whales were calculated in real-time, and locations determined by crossing two or more bearings. A total of 4331 km of ''on-effort'' visual survey resulted in 9 whale sightings compared to 74 acoustic detections. Acoustic detections from 350 h of monitoring resulted in an abundance estimate of 116 (95% CI: 72 293) whales in February, and 123 (95% CI: 77 313) whales in March. The paucity of visual sightings is attributed to the prevailing high winds and the brief periods that whales spend at the surface. These results demonstrate the advantage of acoustic survey methods over visual methods for detecting whales in areas with poor visibility conditions such as the Eastern Caribbean.

8:50

3aABa5. Acoustic and visual monitoring for marine mammals at Cortez and Tanner Banks. Erin M. Oleson, John A. Hildebrand (Univ. of California San Diego, Scripps Inst. of Oceanogr., La Jolla, CA 92093), Mark A. McDonald (Whale Acoust., Laramie, WY), and John Calambokidis (Cascadia Res. Collective, Olympia, WA)

We have implemented continuous acoustic monitoring for baleen whales in the region surrounding Cortez and Tanner Banks, offshore of Southern California, complemented by bimonthly ship-based visual and acoustic observations. Our objective is to develop methods for acoustic monitoring of baleen whales using seafloor acoustic recording packages, the results of which will be used to produce abundance estimates. The joint acoustic and visual effort will answer questions regarding the probability for acoustic detections of individual animals, the quantitative association of calls with a particular species, and how well the number of animals in an area can be assessed acoustically. These findings can then be used to generate a detectability curve for each baleen whale species, from which the seafloor recorder data can be translated into absolute abundance estimates. The use of continuous acoustic monitoring with seafloor recording packages offers the potential for efficient and economical monitoring of marine mammals. The continuous acoustic census approach will be discussed, and compared to opportunistic visual and acoustic survey methods.

9:05

3aABa6. Simple methods for locating, counting, and tracking sperm whales underwater in three dimensions. Duncan E. McGehee (BAE Systems, 4669 Murphy Canyon Rd., Ste. 102, San Diego, CA 92123, duncan.mcgehee@baesystems.com) and John A. Hildebrand (Scripps Inst. of Oceanogr., La Jolla, CA 92037-0205)

Sperm whales' clicks can be used to locate source animals in three dimensions using volumetric arrays of four or more sonobuoys. Clicks are detected using algorithms that depend on gender. Males are easier to track: their clicks are loud and distinctive, and there is usually only one male nearby. A simple envelope detector suffices. Females click more quietly, at a higher rate, and they often occur in large groups, resulting in a continuous barrage of clicks. Using a single click as a matched filter detector enhances the detection of clicks coming from a particular animal. Arrival time differences between the different sonobuoys are used to locate the source in three dimensions. A particular animal can be tracked by determining its location over time. An estimate of the number of phonating animals can be made by locating all the source animals during a particular time interval. During NOAAs SWAPS97 sperm whale survey, arrays of four sonobuoys each were placed around groups of diving sperm whales. Clicks from the animals were recorded at 48ksamples/s. Analyses of these data will be presented. [Work sponsored by NOAA and ONR.]

9:20

3aABa7. Three-dimensional localization of diving sperm whales using a short-aperture towed horizontal array. Aaron M. Thode (MIT, 77 Massachusetts Ave., Cambridge, MA 02139), David K. Mellinger (Oregon State Univ., Newport, OR 97365), Sarah Stienessen (North Seattle, WA 98103), Anthony Martinez (Natl. Marine Fisheries Service, Key Biscayne, FL 33149), and Keith D. Mullin (Natl. Marine Fisheries Service, Pascagoula, MS 39568)

Three-dimensional sperm whale localizations are typically obtained via recording their ''clicks'' on widely distributed hydrophones. Here, an alternative 3D localization method is investigated, using data collected by an 8-m aperture, five-element horizontal towed array deployed from the NOAA ship GORDON GUNTER in the Gulf of Mexico, as part of a recent sperm whale pilot study. During the night of 3 July 2000, the GUNTER towed the array at a steady 1-kn speed and 60-m depth through a pod of sperm whales, in 1000-m-deep water. During that time, surface and bottom reflections from a single sperm whale click were often recorded. By measuring the bearings and relative arrival times of the direct arrival and reflections, the whale location, array depth, and array tilt could be computed. The latter results were checked against measurements obtained from a time-depth recorder attached to the array. Assuming that the speed of the ship remained fixed and the prevailing currents did not change, subsequent positions could then be obtained using only a surface reflection. This method may provide a convenient way for observing dive profiles of animals within a few kilometers of a survey ship. [Work supported by Minerals Management Service, National Marine Fisheries Service, and ONR.]

9:35

3aABa8. Comparison of acoustic and visual surveying of humpback whales off East Australia. Michael J. Noad (Dept. of Veterinary Anatomy and Pathol., Univ. of Sydney, NSW 2006, Australia) and Douglas H. Cato (Defence Sci. and Technol. Organisation, Pyrmont, NSW 2009, Australia)

Many species of whales produce intense sounds that are audible to substantial distances and thus may be useful in censusing, especially in conditions where visual methods have limited effectiveness. Acoustic methods of surveying, however, have their own limitations and are relatively untried compared with visual surveying. This paper describes the results of an experiment to test the effectiveness of acoustic surveying, by applying it to the East Australian humpback whale population, which has been well-surveyed visually. At the experimental site in southeast Queensland, most of this stock passes within visual range of the coast during migrations. Whales were tracked visually using a theodolite on a hill near shore and acoustically using three buoy systems which received the sounds and transmitted the signals to shore. Almost all whales within 10 km could be tracked visually, thus providing a benchmark for the acoustic surveying. Although only a proportion of the stock is vocalizing at any time, the result tests how representative this is of the total stock. The results show that acoustic surveying can be useful for censusing.

9:50

3aABa9. Planning for a pilot census of marine life in the Gulf of Maine: The role of acoustics. Kenneth G. Foote (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, kfoote@whoi.edu)

Preparations are being made for a pilot census of marine life in the Gulf of Maine ecosystem. The role of acoustics as a rapid, remote sensing tool is elaborated. Potential target organisms for acoustic surveying range from mesozooplankton and macrozooplankton to fish and cetaceans. A number of methodological problems must be addressed. These are illustrated for the echo integration method as applied to a stock of Atlantic herring (*Clupea harengus*) [J. Acoust. Soc. Am. **105**, 995 (1999)]. Particular problems of determining target strength and compensating for possible behavioral effects are also general to the method. [Work supported by the Alfred P. Sloan Foundation.]

10:05

3aABa10. Localizing marine animals and how marine animals might localize sound. Gerald L. D'Spain and Paul A. Lepper (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0704)

We can locate vocalizing marine animals, and the animals themselves might locate sound sources, in one of several ways. Time-of-arrival (phase) differences and amplitude differences of a single arrival between spatially separated ''ears'' are well-known techniques. However, they become ineffective in multipath environments and as the frequency of the sound and/or spatial separation between ears decreases. Another approach is to sense properties of the acoustic field in addition to acoustic pressure. This approach, based on a simple Taylor series expansion of the field and ''$f=ma$,'' apparently is exploited by fish, making them biological equivalents of DIFAR sonobuoys. However, do they also measure acoustic strain rate? An additional method, used by humans, is to take advantage of time-of-arrival (phase) differences between multipath arrivals. The information on source location contained in the resulting interference patterns can be understood in terms of waveguide invariants. These speculations will be illustrated with numerical simulations and actual ocean acousic data. [Work supported by ONR.]

10:20

3aABa11. Acoustic identification of female Steller sea lions. Gregory S. Campbell (Cetacean Behavior Lab., San Diego State Univ., San Diego, CA 92182-4611), Robert Gisner (Code 342, Office of Naval Res., Arlington, VA 22217), and David A. Helweg (Code D351, SPAWARSYSCEN San Diego, San Diego, CA 92152)

Steller sea lions (*Eumetopias jubatus*) breed and rear their young in coastal rookeries dispersed along the northern tier of the North Pacific. Population densities are high and no allomaternal behavior occurs, providing strong selection pressure for mother–pup recognition processes. Mothers and pups establish and maintain contact with individually distinctive vocalizations. Our objective is to understand the acoustic features that serve to identify individual females, and develop a ruggedized computer system to perform acoustic recognition of females in the field. We have cataloged almost 2000 contact calls from 46 females in 1998 and 25 in 1999. Each female is visually identified by marking patterns, which provides the ground truth for acoustic identification. Acoustic properties of the calls were measured and presented to several statistical classifiers. Representations of the calls had to be robust with respect to acoustical variability introduced by motivational changes as the mother and pup regained proximity. The calls, classifiers, and results of generalization tests will be described, and a concept for the field system will be discussed. [Research supported by ONR Research Opportunities for Program Officers award N00014-00-1-0114 to R. H. DeFran, Cetacean Behavior Laboratory, SDSU.]

WEDNESDAY MORNING, 6 DECEMBER 2000 SCHOONER/SLOOP ROOMS, 10:50 TO 11:55 A.M.

Session 3aABb

Animal Bioacoustics: General Topics in Bioacoustics

Lawrence F. Wolski, Chair
Hubbs-Sea World Research Institute, 2595 Ingraham Street, San Diego, California 92109

Chair's Introduction—10:50

Contributed Papers

10:55

3aABb1. Infrasonic and low-frequency vocalizations from Siberian and Bengal tigers. Elizabeth von Muggenthaler (Fauna Communications Res. Inst.)

Tigers have many vocalizations including chuffling, growling, prusten, gurgling, grunting, and roaring. It has been well documented that the tiger's high-amplitude, low-frequency roars, which are thought to be territorial in nature [C. Packer and A. E. Pusey, Sci. Am. **276**, 52–59 (1997)] transmit for miles. It has been suggested that because some tigers inhabit dense jungles with limited visiblity, the capacity to hear low frequency may be beneficial for sensing and locating prey [G. T. Huang, J. J. Rosowski, and W. T. Peake, J. Comp. Physiol. A (2000)]. In an effort to understand more about these low-frequency vocalizations and to provide data to other researchers testing hearing in anesthetized felids, 22 tigers, both Siberian and Bengal, are being recorded. A portable system can record from 3 Hz to 22 kHz. On-site real-time analysis of vocalizations is performed using a portable computer. Real-time and edited playback of sonic and infrasonic tiger vocalizations is facilitated by car audio speakers capable of producing frequencies from 10 Hz–22 kHz. Initial findings have documented fundamental frequencies of some roars at 17.50 Hz. Other vocalizations, including chuffling, have fundamental frequencies of 35 Hz ± 5. Playback of both real-time and edited vocalizations appear to illicit behavioral responses, such as roaring, from male tigers.

11:10

3aABb2. On the sound of snapping shrimp: The collapse of a cavitation bubble. Michel Versluis, Anna von der Heydt,[a] Detlef Lohse (Dept. of Appl. Phys. and J. M. Burgers Res. Ctr. for Fluid Dynam., Univ. of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands), and Barbara Schmitz (TU Munchen, 85747 Garching, Germany)

Snapping shrimp produce a snapping sound by an extremely rapid closure of their snapper claw. They usually occur in large numbers providing a permanent crackling background noise, thereby severely limiting the use of underwater acoustics for active and passive sonar, both in scientific and naval applications. Source levels reported for *Alpheus hetero-*

chaelis are as high as 220 dB (peak-to-peak) $re\ 1\ \mu$Pa at 1 m distance. Recent ultra-high-speed imaging of the snapper claw closure [Versluis *et al.*, Science (in press)] revealed that the sound is generated by the collapse of a cavitation bubble formed in a fast flowing jet of water forced out from between the claws during claw closure. In this work, we develop a theoretical model for a bubble under such conditions. The dynamics of the bubble radius and the emitted sound can be described by the Rayleigh–Plesset equation. The calculated results are compared with the experimental data. The model fully reproduces the bubble dynamics and it quantitatively accounts for the time dependence of the bubble radius and for the emitted sound. [a] Also at Dept. of Physics, Philipps-Univ. Marburg, Renthof 6, 35032 Marburg, Germany.

11:25

3aABb3. A unique way of sound production in the snapping shrimp (*Alpheus heterochaelis*). Barbara Schmitz (Dept. of Zoology, TU Muenchen, Lichtenbergstr. 4, 85747 Garching, Germany), Michel Versluis, Anna von der Heydt, and Detlef Lohse (Appl. Phys., Univ. of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands)

Sound production is known in more than 50, mostly stridulating, crustacean genera. These acoustic signals occur in agonistic interactions as well as for mate attraction. The mechanism of sound production in snapping shrimp, which also serves to stun or even kill small prey, is especially interesting. The current assumption was that the sound is produced by cocking and then rapidly closing the enlarged modified snapper claw. Snapping shrimp sounds contribute most to coastal biological noise, may be heard up to 1 mile away, and resemble the crackling of dry twigs in fire or the sizzle of frying fat. Recent hydrophone measurements close to tethered shrimp (*Alpheus heterochaelis*) revealed pulse-like signals of 500-ns duration, comprising frequencies beyond 200 kHz, and showing enormous sound pressure levels of up to 220-dB $re\ 1\ \mu$Pa (peak to peak) at 1-m distance. Such high intensities are very unlikely to be produced by the mechanical contact of two claw surfaces. Ultra-high-speed video recordings and simultaneous hydrophone measurements reveal that claw closure results in a water jet, the high velocity of which (25 m/s) leads to the formation of a cavitation bubble, which emits the extremely loud sound upon its collapse.

11:40

3aABb4. Lack of species-specific vocal recognition in Amazonian manatees: *Trichechus inunguis*. Renata S. Sousa Lima and Vera M. F. da Silva (Laboratório de Mamíferos Aquáticos, INPA, C.P. 478, Manaus, AM, Brasil 69083-000, pboi@inpa.gov.br)

Playback experiments were conducted in order to test for the existence of species-specific vocal recognition in Amazonian manatees. The animals were isolated in pools while acoustic stimuli were played from a tape recorder and transmitted underwater through a loudspeaker. Nine animals were monitored for response to playback vocalizations from eighteen different individuals, nine from each species (*Trichechus inunguis* and *T. manatus manatus*). No significant differences were detected in the response of manatees exposed to the different stimuli. Only the time spent close to the speaker was greater when the animals were exposed to conspecific vocalizations ($\rho = 0.375$). This result suggests that Amazonian manatees cannot recognize differences between their own and another manatee species' vocalization. The methodology was also tested and no difference in the response of animals exposed to silence or to tape hiss (blank tape used as control) was found. Testing the response to the presence or absence of vocalizations, significant differences were found in time elapsed between the playback and the response ($\rho < 0.001$) and in the time spent next to the speaker ($\rho = 0.032$), confirming their great ability to perceive sounds underwater. [Work supported by FBPN, MacArthur Foundation, CI, FINEP, MCT/PPG7 and CNPq.]

WEDNESDAY MORNING, 6 DECEMBER 2000 CALIFORNIA SALON 3, 7:20 A.M. TO 12:05 P.M.

Session 3aAO

Acoustical Oceanography: Special Topics

Thomas K. Berger, Cochair
Scripps Institute of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0238

Christopher D. Jones, Cochair
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

Chair's Introduction—7:20

Contributed Papers

7:25

3aAO1. The *F* factor: Ocean climatology and internal-wave acoustic effects. Stanley M. Flatté and Kimberly J. Noble (Phys. Dept. and Inst. of Marine Sci., Univ. of California at Santa Cruz, Santa Cruz, CA 95064)

The strength of acoustic fluctuations due to internal waves is affected by the temperature-salinity relation of an oceanic region. We define an acoustic fluctuation strength parameter F as the ratio of the fractional potential-sound-speed change to the fractional potential-density change. F is calculated at three depth levels (275, 550, and 850 m), on a one-degree grid of latitude and longitude, using NODC/OCL's World Ocean Atlas 1994. Representative values of F in upper waters range between 5 and 35. Results for intermediate depths range from 5 to 60. In general, F exhibits higher values in the Atlantic Basin than in the Indian or Pacific, and has a maximum at 550 m. The main use of F will be the prediction of travel-time fluctuations in acoustic propagation experiments, which will be proportional to the value of F, given a universal strength of internal waves.

7:40

3aAO2. Imaging acoustic fluctuations in shallow water using dislocation theory. D. P. Williams, G. L. D'Spain, W. S. Hodgkiss, and W. A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., Univ. of California at San Diego, La Jolla, CA 92093-0238)

The distribution of the amplitude of a moderate-frequency sound field in a shallow ocean is considered in relation to the existence of dislocations in the phase front [Nye and Berry (1974)] where the amplitude is close to zero. Phase front dislocations strictly occur when the amplitude is zero; in actual ocean acoustic measurements, only low-field amplitudes at a minima can be distinguished due to the interference of background noise.

The dislocations are sensitive to temporal variability in the ocean environment [Kuz'kin, Ogurtsov, and Petnikov (1998)] and their space-frequency fluctuations lend qualitative insight into environmental characteristics. A model is developed to predict the patterns of the sound-field minima resulting, in particular, from internal wave-induced fluctuations in the environment. The model is then applied to the interpretation of the results of 2–7-kHz broadband transmissions collected during a noisy shallow-water acoustics experiment off the coast of San Diego. [Work supported by ONR.]

7:55

3aAO3. Energetics of dislocations in ocean acoustic fields. Gerald L. D'Spain, Duncan P. Williams, and William A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0704)

Dislocations are places in the acoustic field where the amplitude goes to zero and the total field phase becomes undetermined [Nye and Berry (1974)]. The positions of these dislocations are sensitive indicators of changes in the ocean environment. Expressions for the acoustic potential and kinetic energy densities and active and reactive vector acoustic intensities are derived for a single-tone sound field. These expressions are then applied to infrasonic ocean acoustic data collected by the Marine Physical Lab's set of freely drifting Swallow floats to help identify dislocations in the underwater acoustic pressure and the component particle velocity fields. Likely dislocation positions were identified during cw source tows at 7, 10, and 16 Hz during the 1990 NATIVE 1 experiment near the Blake Plateau in the northwest Atlantic Ocean. The neighborhoods of dislocations, as well as near other amplitude minima, are regions where a significant acoustic energy flow occurs. The relationship between dislocations and polarization of the acoustic particle motion also is discussed. [Work supported by ONR.]

8:10

3aAO4. Fourier component fluctuations of wideband ocean tomography signals: Empirical temporal statistics from AST96. Rex K. Andrew, Bruce M. Howe, and James A. Mercer (Appl. Phys. Lab., 1013 NE 40th St., Seattle, WA 98105, randrew@apl.washington.edu)

Current theories for wave propagation in random media can predict fluctuations for narrow-band (single-frequency) signals, but do not explicitly address wideband signals. Recent ocean acoustic tomography experiments, however, have employed wideband ''m-sequence'' signals. Processing here assumes a periodic signal, and hence, involves a unique set of Fourier series components. A technique is presented for decomposing such wideband signals into their Fourier components and then analyzing the space–time fluctuations of these ''narrow-band'' components. The technique is applied to AST96 data transmitted over 150-km and 1-Mm paths for single-point and two-point (temporal separation) statistics. [Work support by ONR.]

8:25

3aAO5. Hilbert spectrum of internal waves. Tao Wang (California Inst. of Technol., Pasadena, CA 91125) and Tong Su (Marine Systems Eng. Res. Inst., China State Shipbuilding Corp., Beijing 100036, PROC)

The newly developed Hilbert spectrum offers a powerful tool for nonlinear, nonstationary data analysis, and has found wide applications to Stokes waves, random ocean wind waves, and turbulence [N. E. Huang, Z. Shen, and S. R. Long, Annu. Rev. Fluid Mech. **31**, 417–457 (1999)]. This paper presents the Hilbert spectrum of internal waves in the Yellow Sea and East China Sea. According to their intrinsic characteristic scales, the internal wave data are first decomposed into a number of modes of different time scales, designated as intrinsic mode function (IMF) components. The Hilbert transform is then applied to the IMF components to construct the frequency–time distribution of energy density, the Hilbert spectrum. It is shown that the Hilbert spectral analysis provides a new physical insight in the underlying dynamic processes of these shallow-water internal wave phenomena.

8:40

3aAO6. Matched field inversion with a moving source in a shallow-water environment. John Viechnicki and Ross Chapman (Ocean Acoust. Group, School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada, jayvee@ducks.seos.uvic.ca)

Geoacoustic inversion based on matched field processing (MFP) is examined for shallow-water, low-frequency environments. Specific interest lies with resolving geoacoustic parameters from cw tones projected from a moving source. Data obtained using vertical line arrays (VLA) are available from both the 1996 Haro Strait PRIMER Experiment (HSX) and the Santa Barbara Channel Experiment (SBCX) of 1998. The environmental complexity associated with these experiments, namely, strong range-dependent bathymetry and current flow, is typical to littoral environments in general and must be appropriately addressed. Parabolic equation modeling is used as it provides range-dependent results. The VLA receiver configuration is described as a catenary which is typical of bottom-anchored arrays drifting in uniform current flow. Both SBCX and HSX are useful for benchmarking geoacoustic inversion techniques since results from other techniques are available. Estimation of bottom properties is discussed as a function of propagation range, ship track with respect to receiver position, and general bathymetric features. Results for both tangential and radial tracks are presented. [Work supported by ONR.]

8:55

3aAO7. The effect of internal wave time dependence on ray chaos. Michael Vera and Stanley M. Flatté (Phys. Dept. and Inst. of Marine Sci., Univ. of California at Santa Cruz, Santa Cruz, CA 95064)

A range-dependent field of sound speed in the ocean can give rise to chaotic instabilities in acoustic ray paths. A model of internal waves is used as a range-dependent effect on the speed of sound in numerical simulations. These simulations, performed over a 1000-km range, demonstrate the sensitivity of ray paths to slight changes in the sound-speed field or the initial conditions of their launch. Previous work [Simmen *et al.* (1999)] has shown the effect on arrival depth of slightly shifting the launch angle of each ray. The focus of this work is the effect of internal wave time dependence. Comparisons are made between simulations using a time-dependent medium and those using a time-independent (''frozen'') medium, and between simulations with different geophysical starting times. The resulting arrival depth differences are compared to those induced by launch angle shifts. The structure and magnitude of these differences are considered in the case of two different sound-speed profiles: a Munk canonical profile and a profile from the Slice89 experiment.

9:10

3aAO8. Characterization of the internal waves and its effect on signal propagation in the Adventure Bank area. T. C. Yang, Kwang Yoo (Naval Res. Lab., Washington, DC 20375), and Martin Siderius (SACLANTCEN, La Spazia, Italy)

During the Advent-99 experiment, which took place in the Sicily Strait in May, 1999, multitone (in the band 200–1500 Hz) narrow-band signals were transmitted from a bottom-mounted tower and received on a 64 phone line array (VLA) covering the water column; the water depth was 80 m. Data at a source–receiver range of 10 km will be presented and modeled using an internal wave model. A CTD chain containing 48 CTDs on a vertical string was towed continuously during the acoustic experiments between the source and VLA, yielding a time-evolving sound speed profile between the source and receiver. From that the average buoyancy frequency and sound speed profile were determined as a function of range. Internal-wave-mode depth functions were calculated using the measured buoyancy profiles. Internal-waves-mode amplitudes were determined from the eigenvalues of the sound speed covariance matrix for the lowest 5 modes. Internal-wave-frequency spectrum has the typical 2-power dependence as in the Garrett–Munk model. The modeled results of mean transmission loss and amplitude fluctuations with and without internal waves will be presented and compared with data. [Work supported by the Office of Naval Research and SACLANTCEN.]

9:25

3aAO9. Sound in a time-dependent ocean: Implications on acoustic tomography of currents. Oleg A. Godin (NOAA/Environ. Technol. Lab., Boulder, CO 80305)

Ocean is a nonstationary acoustic environment. Surface and internal waves, tides, and mesoscale eddies contribute to sound-speed and ocean surface variability on different temporal scales. Acoustic effects of the nonstationarity include violation of the reciprocity principle and signal frequency variation along the propagation path. Mathematical models of underwater sound propagation often either ignore ocean nonstationarity or account for it within the frozen-medium approximation. In this paper, accuracy of different approaches to model sound propagation in time-dependent ocean is analyzed within the frameworks of the ray theory and the adiabatic mode approximation. The theoretical approach is based on the space-time geometrical acoustics and a space-time version of the ''vertical modes–horizontal rays'' technique. While the frozen-medium approximation proves to be generally rather crude and inadequate in modeling nonreciprocity due to sound-speed time dependence, another simple technique, the quasistationary approximation, is shown to be a sufficiently accurate and efficient approach to modeling low-frequency underwater sound propagation. Contributions to ray travel time and mode-phase nonreciprocity due to medium motion and time dependence are compared for several typical current tomography scenarios. The feasibility of distinguishing between acoustic nonreciprocities due to currents and due to medium nonstationarity is discussed. [Work supported by NRC.]

9:40–9:50 Break

9:50

3aAO10. Frequency dependent transmission loss related to modal properties in layered nonlinear internal wave environments. Scott Frank, William L. Siegmann (Dept. of Math. Sci., Rennselaer Polytechnic Inst., Troy, NY 12180), Mohsen Badiey, Yongke Mu (Univ. of Delaware, Newark, DE 19716), and James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Broadband data were taken during the SWARM 95 experiment from airgun and linear frequency modulated sources while nonlinear internal waves were crossing two different propagation tracks. In a previous study, the geotime behavior of the depth-averaged received intensity was examined, and amplitude modulations were associated with the passage of internal waves. Additionally, a time-frequency analysis of the data exhibited periodic changes in the modal structure of the signals. Though some discrepancies occurred, both of these effects were observed in parabolic equation simulations of the SWARM environment. To address the discrepancies in received intensity levels, various physical mechanisms will be investigated, for example geoacoustic attenuation and ocean thermocline variability. Effects on the time-frequency representations of signals passing through internal wave packets will be documented and categorized. To understand features of frequency dependent transmission loss, relationships between the eigenvalues and the variable layer structure of the waveguide will be developed and applied to both measurements and simulations. Such relationships will aid in extracting the amplitude, width, and azimuthal propagation angle of nonlinear internal waves from broadband acoustic data. [Work supported by ONR.]

10:05

3aAO11. One-path scintillation method for transverse current monitoring. Iosif Fuks, Mikhail Charnotskii, and Konstantin Naugolnykh (Univ. of Colorado, CIRES/NOAA, Environ. Technol. Lab., 325 Broadway, Boulder, CO 80305)

A sound wave passing through the inhomogeneous fluid is modulated producing the intensity fluctuations or scintillations at the receiving plane. These fluctuations evolve as the intervening medium changes. If the fluctuations in the medium are produced by the advection of frozen fine-structure field, then the evolution of the scintillation pattern can be linked to the motion of the ocean medium [S. Clifford and D. Farmer, J. Acoust. Soc. Am. **74**, 1826–1832 (1983)]. This method of the ocean current registration can be realized by the measurement of correlation of the signals transmitted from the source to two separated in space receivers and can be respectively considered as space-domain scintillation method. However, it is possible to register the transverse flow by sounding ocean along the one path using the multifrequency signal [J. Fuks, Radiotechn. Electron. (in Russian) **3**, 515–524 (1975)]. The coherence of frequency-spaced signal fluctuation is related to the advection of the inhomogeneous medium through the sound path, thus providing a basis for the current velocity measurement. The sensitivity of this method in application to flows in straits and narrow channels is considered and the accuracy of the current velocity measurement is estimated.

10:20

3aAO12. Beamforming through an anisotropic, time-varying internal wave field. Steven Finette and Roger M. Oba (Naval Res. Lab., Washington, DC 20375-5350)

Shallow-water environments commonly exhibit complex spatial and temporal variations in temperature and salinity caused by internal wave activity. These quantities map into a sound-speed distribution that can significantly distort both the amplitude and phase of an acoustic field propagating through the environment. This presentation describes computer simulation results on the effect of an anisotropic, time-varying internal wave field on plane wave beamforming of a point source within a shallow-water waveguide. The (azimuthal) anisotropy is described by a solitary wave packet propagating through the region. A 3D parabolic equation code was used to propagate a 400-Hz signal through this space and time-evolving environment. Beamformed power, parametrized by receiver depth and an azimuthal angle, is illustrated for horizontal arrays and presented in terms of beam space, time, and range from the source. When the packet enters the region between source and receiving array, significant periodic beam wander and beam splitting are seen; such effects are shown to be related to horizontal refraction of energy from individual internal wave depressions forming the solitary wave packet. [Work supported by the Office of Naval Research.]

10:35

3aAO13. Acoustic detection and classification of hurricanes. Joshua D. Wilson and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

Hurricanes are some of the most destructive natural disasters known to man. Currently, it is difficult to detect, track, and classify hurricanes in areas of the world that lack satellite coverage and specialized aircraft. The purpose of this research is to investigate the feasibility of using underwater hydrophone arrays to detect and classify a tropical cyclone. A wave theoretic model is developed for the spatial covariance of the ambient noise field generated by a hurricane. Simulations are performed to determine whether ambient noise due to a hurricane is detectable over the more typical wind-wave noise accumulated from the rest of the ocean. Scenarios where a hydrophone array is placed directly beneath the storm and at a remote distance from the storm are investigated.

10:50

3aAO14. Acoustic observation of hydrothermal flows. Christopher D. Jones, Darrell R. Jackson (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105), Peter A. Rona, and Karen G. Bemis (Inst. of Marine and Coastal Sci., Rutgers Univ., New Brunswick, NJ 08901-8521)

Acoustic backscatter has been used to observe hydrothermal flows at several sites on the Juan de Fuca Ridge at frequencies of 200 and 330 kHz. In the case of plume discharges (''smokers''), backscatter from particulates is used to obtain three-dimensional images and to measure flow via Doppler shift. For diffuse flows, acoustic scintillation of the signal backscattered from the seafloor is detected using ping-to-ping correlation. Mapping of diffuse flows using both fixed and moving platforms will be

discussed. The magnitude and time lag dependence of ping-to-ping correlation have been linked to flow parameters using the Rytov scattering approximation. [Work supported by NSF Grant OCE 98-18841.]

11:05

3aAO15. Low-frequency acoustics of bubble plumes formed in fresh water and salt water. Thomas K. Berger and Michael J. Buckingham (Scripps Inst. of Oceanogr., Univ. of California San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0238, tberger@ucsd.edu)

Experiments aimed at comparing the acoustics of bubble plumes formed in salt water and fresh water were performed in a laboratory tank and natural bodies of water. The plumes were generated as a vertical stream of water, with a diameter of millimeter order and velocity ranging to about 10 m/s, penetrated a free water surface. The acoustic signals were measured with hydrophones in different positions, amplified, and digitally sampled. The measured power spectra revealed well-defined, nonuniformly spaced resonance peaks at frequencies below about 1 kHz. A theory based on a conical bubble plume, with a nonuniform sound speed profile acting as a resonant cavity, predicts the frequencies of the resonance peaks and shows remarkable agreement with the experimentally measured values for both salt water and fresh water bubble plumes. Though the bubbles contained in the salt water bubble plumes are much smaller than those in the fresh water plumes, the resonance frequencies are largely unaffected, consistent with the observation that the plumes in both cases are nearly identical in size and geometry for the same incoming velocity and air-entrainment rate. The acoustic levels in salt water were lower, indicating possible differences in attenuation or driving mechanisms. [Work supported by ONR.]

11:20

3aAO16. Acoustic propagation measurements in the surf. John S. Stroud, Kerry W. Commander (Coastal Systems Station, Code R21, 6703 W. Hwy. 98, Panama City, FL 32407-7001, stroudjs@ncsc.navy.mil), Robert J. McDonald, and Jo Ellen Wilbur (Systems Station, Panama City, FL 32407-7001)

As a follow-on experiment to the 1999 Near Shore Acoustic Network Experiment, the 2000 Surf Zone Acoustic Test Experiment (SZATE) was conducted off of the Scripps Institute of Oceanography pier in La Jolla, CA. In this experiment, measurements were conducted in a very shallow water/surf zone (VSW/SZ) region to investigate acoustic propagation and sound channel stability in the coastal area. These measurements utilized two receivers and a broad band source. The data to be presented were acquired by a computer controlled system designed and operated by Grant Deane of Scripps Institution of Oceanography. This system provided for 12 18-min data acquisition periods per day. The probative transmissions consisted of a variety of linear frequency modulated (LFM) and binary phase shift keyed (BPSK), as well as 4-ary frequency shift keyed (FSK) signals. Each of the types of signals was used during each data acquisition period, providing a record of propagation throughout the tidal cycle. Results of this acoustic measurement for select pulse types will be reported. [Work supported by ONR Code 321OE.]

11:35

3aAO17. Monostatic and bistatic scattering by a single bubble near a pressure release interface: Laboratory measurements and modeling. George Kapodistrias and Peter H. Dahl (Appl. Phys. Lab. and Dept. of Mech. Eng., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, georgek@apl.washington.edu)

Scattering by a single bubble near a flat pressure release interface is investigated theoretically and experimentally. A ray-acoustic interpretation is used to describe the four scattering paths, from source to bubble to receiver. Multiple scattering effects are accounted for using a closed-form solution derived from the multiple scattering series. A bubble with radius $a \approx 425$ μm is placed on a fine nylon thread, which is practically transparent to sound, at a distance d_B from the interface. The primary variable is d_B and it ranges from $1a$ to $100a$. Experiments are performed at a frequency of 120 kHz with the transducers arranged in both bistatic and monostatic configuration. Theory and experiment are in excellent agreement, verifying the dominant effect of the four paths in the response of the bubble, with multiple scattering playing a role for $kd_B < 1$, where k is the wave number of the medium. In the long-range limit our simulations agree with those of Ye and Feuillade [J. Acoust. Soc. Am. **102**, 798–805 (1997)] including the shifting of the bubble's resonant frequency. The dependence of scattering on transducer arrangement, range to bubble, grazing angle, and phase relation among the four paths, *vis-à-vis* monostatic and bistatic scattering, are discussed.

11:50

3aAO18. Space-time sound fluctuations due to internal solitons in shallow water. Boris G. Katsnelson, Serguey A. Pereselkov (Voronezh Univ., 1 Universitetskaya Sq., Voronezh 394693, Russia), Valery G. Petnikov (General Phys. Inst., Moscow 117333, Russia), Konstantin D. Sabinin, and Andrey N. Serebryany (N. N. Andreev Acoust. Inst., Moscow 117333, Russia)

In this paper results of numerical modeling of sound propagation in shallow-water waveguide with internal waves are presented on the basis of experimental data about internal solitons (IS) registered in the Japan sea [A. N. Serebryany, Sov. Phys. Oceanol. **27**, 225–226 (1987)]. These IS constitute some sequences of oscillations with amplitude up to 10 m moving toward coastal line with the velocity about 0.25 m/s. The wavelength of separate oscillations in this group is about 325 m; total length of the packet is about 4 km. Three-dimensional sound-speed distribution is constructed using the mentioned data. On the basis of this model, space-time fluctuations of intensity of the sound propagating in such medium are calculated. The modeling is carried out within the framework of the parabolic equation method in the horizontal plane and the modal approach in the vertical one. These fluctuations can be interpreted as the forming of a sound waveguide in the horizontal plane; space-time scales of fluctuations correlate with space-time scales of IS. Different numerical examples are presented; possible experimental setup is discussed. [Work supported by RFBR, Grant 00-05-64752.]

Session 3aBB

Biomedical Ultrasound/Bioresponse to Vibration: Therapeutic and Diagnostic Ultrasound

Mark R. Prausnitz, Chair
School of Chemical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0100

Chair's Introduction—7:30

Contributed Papers

7:35

3aBB1. Real-time observation of inception and growth of HIFU-induced tissue lesions. Cyril Lafon, Michael R. Bailey, Lisa N. Couret, Peter J. Kaczkowski, Andrew A. Brayman, Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105), and Oleg A. Sapozhnikov (Moscow State Univ., Moscow, Russia)

To study the biological effects of high-intensity focused ultrasound (HIFU), experiments are usually performed on isolated or perfused tissues. Indeed, the complex phenomena occurring in tissue during HIFU-induced coagulation necrosis is difficult to mimic with synthetic phantoms. A good phantom should first match the acoustical and thermal properties of tissues. Furthermore, heating above a thermal threshold should induce a permanent, localized and observable change corresponding to protein denaturing in tissue. Lastly, the choice of a transparent material makes possible real-time examination of the development of coagulation necroses. We have used bovine eye lenses in this aim. The density, sound speed, attenuation, and thermal threshold for irreversible damage to the bovine lens were measured and found to be similar to those for liver or muscle, common tissues for HIFU experiments, although acoustic attenuation is slightly higher in the lens. Transparency of the lens allowed us to observe HIFU-induced lesion evolution in real time. The shape and size of the lesions obtained in the lens agreed well with results obtained in liver. In conclusion, the transparent bovine eye lens is a useful model for visualization of thermal lesions. [Work supported by ONR.]

7:50

3aBB2. Theoretical predictions of ultrasonic fields, temperature response, and lesion dynamics in biological tissue for the purpose of noninvasive disease treatment. Francesco P. Curra, Pierre D. Mourad, Steven G. Kargl, Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105), and Vera A. Khokhlova (Moscow State Univ., Moscow, Russia)

Ultrasound has been used for decades as a means for noninvasive treatment of diseases. Low-intensity ultrasound is routinely applied in physical therapy for muscular and neurological related illnesses. In contrast, high-intensity focused ultrasound (HIFU) is used to induce coagulative necrosis of tissue for cancer treatment or hemostasis. Our efforts concern the latter. Predictions of ultrasound fields, temperature response, and lesion dynamics are obtained by a model which accounts for nonlinear sound propagation in inhomogeneous media, an arbitrary frequency power law for acoustic attenuation, and temperature time history [J. Acoust. Soc. Am. **107**, No. 5, Pt. 2 (2000)]. The model is expanded from its previous version to include attenuation and sound speed dependence on temperature levels and also to consider generation of gas bubbles within the tissue. Results are presented in terms of treatment strategies that provide maximum energy transfer for coagulating the targeted tissue while minimizing damage to the surrounding area.

8:05

3aBB3. Two-dimensional ultrasound phased array for treatment of benign prostatic hyperplasia. Janelle L. Helser, Victor W. Sparrow, and Nadine Smith (Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804, nbs@engr.psu.edu)

In the prostate, focused ultrasound offers an attractive means of noninvasive tissue ablation for treatment of Benign Prostatic Hyperplasia (BPH). Intracavitary two-dimensional, ultrasound arrays allow for deep localized heating and are capable of generating sufficient power for tissue ablation. A two-dimensional array design is capable of focusing electronically in three dimensions, by controlling the power and phase, instead of having to physically adjust or reposition the transducer. The goal of this project was to design a 1.5 MHz, two-dimensional, linear, PZT transducer that could steer a focus throughout the region of the prostate. Pressure fields for the array were theoretically determined, using amplitude shading and varying elemental sizes to reduce grating lobe and side lobe levels. The amplitude shading was altered each time the transducer was refocused so that the peak amplitude was located in the element normal to the focus. Based on this 100-element theoretical array, a 1.5 cm$\times$8 cm transducer array was constructed to evaluate the feasibility of delivering thermal therapies from a two-dimensional array.

8:20

3aBB4. High-intensity focused ultrasound hyperthermia in nonuniform tissue phantoms. Jinlan Huang, R. Glynn Holt, and Ronald A. Roy (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215, jinlan@bu.edu)

It has been demonstrated that high-intensity focused ultrasound (HIFU) shows promising potential for stopping bleeding, both from individual blood vessels as well as from gross damage to the capillary bed. Thermal effects have been suggested to play a major role in occlusion of small vessels and also appear to contribute to the hemostasis of major blood vessels. The goal of the present work is to understand the heating process in nonuniform flow-through tissue phantoms. The experiments have been focused on a single-vessel flow-through phantom, and the studies include the prediction and measurement of temperature fields as a function of acoustic pressure, insonation time and flow rate, and investigating different insonation geometries including on-axis and off-axis insonations. The experimental data are compared to simulation results using the Westervelt equation for calculating the nonlinear acoustic field and two coupled bioheat equations for the tissue domain and the blood domain for the temperature field calculation. [Work supported by DARPA.]

8:35

3aBB5. Platelet activity as a result of exposure to high-intensity focused ultrasound. Sandra L. Poliachik, Ryan J. Ollos, Pierre D. Mourad, Lawrence A. Crum (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, poliachi@u.washington.edu), and Wayne L. Chandler (Univ. of Washington, Seattle, WA 98195)

Using platelet-rich plasma, we investigated the capability of 1.1-MHz cw high-intensity focused ultrasound (HIFU) to produce ''acoustic primary hemostasis,'' including platelet activation, aggregation, and adhesion to a collagen-coated surface. Platelet activity was evaluated for exposure

durations of 100–500 s at intensities of 0–2250 W/cm^2. In order to avoid heating effects, temperatures in platelet trials were maintained below 42 °C through use of a tank cooling system and control of exposure parameters. Flow cytometry, laser aggregometry, conventional microscopy, environmental scanning electron microscopy, and passive cavitation detection were used to quantify platelet activation, aggregation, adhesion, and associated cavitation. HIFU can activate platelets and cause them to adhere to a collagen-coated surface. Cavitation was monitored during aggregation trials and was quantified to provide a relative measure of the amount of cavitation that occurred in each aggregation trial. Regression analysis shows weak correlation between aggregation and intensity, and a strong correlation between aggregation and cavitation occurrence. [Research supported by DARPA.]

8:50

3aBB6. Influence of chemical composition of membrane-disrupting polymers on relative cavitation activity and hemolysis. Tyrone M. Porter, Josh Nickerson, Lawrence A. Crum (Appl. Phys. Lab., 1013 NE 40th St., Seattle, WA 98105, tporter@apl.washington.edu), Fiona E. Black, Niren Murthy, Patrick S. Stayton, and Allan S. Hoffman (Univ. of Washington, Seattle, WA 98195)

In previous studies, we have shown that membrane-disrupting polymers may also act as cavitation promoters. These polymers are designed to efficiently disrupt red blood cells in a *p*H-dependent fashion (Murthy *et al.*). When combined with high-intensity focused ultrasound (HIFU), there is a noted increase in relative cavitation activity and a corresponding increase in red blood cell lysis. Varying the chemical composition of the polymer (length of hydrocarbon chains, molecular weight, etc.) modifies the hemolytic and cavitation-promoting activity of the polymer. For example, the hemolytic and cavitation promoting activity of poly(ethyl acrylic acid) (PEAAc) rises as the *p*H of the solvent decreases from 7.4 to 6.1. However, the polymer poly(propyl acrylic acid) (PPAAc), which has a more hydrophobic pendant alkyl group, promotes cavitation and, therefore, hemolysis at a *p*H of 5.0, 6.1, and 7.4. Variations in the polymer molecular weight change the number of hydrophobic regions which also alters the hemolytic and cavitation activity. From these results, polymers may be designed which, when combined with ultrasound, optimize drug transport across cell membranes in a *p*H-dependent or -independent manner.

9:05

3aBB7. Behavior of ultrasound contrast agents near the fragmentation threshold. Wen-Shiang Chen, Thomas J. Matula, and Lawrence A. Crum (Appl. Phys. Lab, Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698, wschen@u.washington.edu)

Understanding the destruction process of ultrasound contrast agents is important in therapeutic ultrasound applications (such as ultrasound-enhanced drug delivery), as well as in certain imaging applications (such as ''flash echo'' imaging of the myocardium). In the destruction of Optison^{TM} microbubbles, our observations suggest that there are two pressure thresholds: a lower threshold which leads to shell rupture and the production of daughter (derivative) bubbles, and a higher threshold leading to the onset of the inertial cavitation (IC) activity. Slightly above the shell-disruption (SD) threshold, the acoustic scattering signal decreases, presumably due in part to the partial dissolution of the derivative bubbles. The extent of the quiescent region, the IC pressure threshold, and the strength of the subsequent cavitation activity are all highly dependent on the acoustic pulse parameters. It is found that for pressure amplitudes in excess of the IC threshold, a longer pulse length or a higher pressure level will decrease the duration of this quiescent period. By controlling the bubbles' response through fine-tuning the acoustic parameters near the SD and IC threshold, one could design an ultrasound system which would optimize the desirable effects of contrast agents for imaging and therapy applications. [Work supported in part by ONR and WTC.]

9:20

3aBB8. Analysis of deformation process for cell model including a gas bubble by shock waves. Masaaki Tamagawa, Ichiroh Yamanoi, and Atsuhi Matsumoto (Grad. School of Energy Sci., Kyoto Univ., Kyoto, 606-8501 Japan)

This paper describes the analysis of deformation process for cell model including a gas bubble by shock waves and flow induced by propagating underwater shock waves, in order to study the mechanism and high efficiency of destruction of cells and microcapsules by shock waves and bubbles, such as for drug delivery system (DDS) and for bioprocesses to make energy and for environmental protection. For this analysis, the cells are modeled as liquid droplet including a gas bubble in water. By computation, using the ALE (arbitrary Lagrangian–Eulerian) method, the coupling oscillation is analyzed for particle velocity and pressure wave through the cell. The results show that the effects of the acoustic impedance of the fluid media in water is large for the deformation process and the effects of deformation of droplet boundary is also large for deformation process for a gas bubble inside the cell and for microjet flow by a collapsing bubble.

9:35

3aBB9. Study of the mechanism of fragmentation of a microbubble exposed to ultrasound using a high-speed observation system. Nobuki Kudo, Takehiro Miyaoka, Kaori Kuribayashi, Katsuyuki Yamamoto (Grad. School of Eng., Hokkaido Univ. N13W8, kita-ku, Sapporo, 060-8628 Japan, kudo@bme.eng.hokudai.ac.jp), and Michiya Natori (Natl. Ohkura Hospital, Tokyo, Japan)

In this study, we observed the dynamic behaviors of microbubbles exposed to ultrasound using a high-speed camera (Ultranac, Nac, Japan). Although the maximum frame rate of the high-speed camera used in this study is 20 million frames/s, the possible frame rate was limited to 6 million frames/s because the high optical magnification of the observation system resulted in a shortage of light exposure to the camera. Behaviors of albumin-shelled microbubbles of 10–60 microns in diameter were clearly observed as in 24-frame sequential high-speed photographs at the maximum optical magnification of 85×. Several typical phenomena of bubble behavior, such as deformation and rupture of the bubble shell, anisotropic contraction of a microbubble and generation of a small stream, and fragmentation of a bubble by the small stream, were observed. From those observation results, it was concluded that the small stream generated by anisotropic contraction of a microbubble exposed to ultrasound causes fragmentation of the bubble in the contraction phase. [This research was partially supported by a grant-in-aid for scientific research from the Ministry of Education, Science, Sports and Culture, Japan.]

9:50

3aBB10. An attempt to explain turbulence occurring when yeast cells suspended in water-rich ethanol mixtures are irradiated in a quasistanding ultrasonic plane wave at 2.2 MHz. Stefan Radel, Aiden J. McLoughlin (Dept. of Industrial Microbiology, Univ. College Dublin, Belfield, Dublin 4, Ireland, radel@iap.twien.ac.at), Martin Groeschl, and Ewald Benes (Inst. of General Phys., Vienna Univ. of Technol., Vienna, Austria)

Ultrasonic resonators are successfully used for the separation of suspended particles or cells, e.g., as cells filters in biotechnology. These systems are based on the radiation forces spatially rearranging the particles in the pressure nodal planes of the quasistanding ultrasonic wave field. When yeast—as this technology is tested to be applied in brewing—is suspended in water this separation/immobilization takes place as predicted. However, suspensions of yeast in water-rich ethanol mixtures show a breakdown of the spatial distribution above certain concentrations of ethanol and the cells are turbulently mixed in the separation system. This has earlier been reported to go along with an impaired viability of yeast in a 12% (v/v) ethanol/water mixture. The scope of this paper is to attempt an explanation based on the acoustic contrast between the cells and the liquid as water-rich ethanol mixtures have unexpected acoustical and physical properties. Data of the spectral behavior of the system at *in situ* levels of true elec-

trical power input are shown as well as the results of examinations, if (sub) harmonic and/or transient cavitation were taking place. [Work supported in part by the European Commission's TMR Programme, Contract No. ERBFMRXCT97-0156, EuroUltraSonoSep.]

10:05–10:15 Break

10:15

3aBB11. Potential errors in contrast-agent-mediated estimates of regional myocardial perfusion arising from myocardial anisotropy. James G. Miller, David E. Sosnovik, Steven L. Baldwin, and Mark R. Holland (Dept. of Phys., Washington Univ., St. Louis, MO 63130)

Recently developed imaging modalities based on the nonlinear response of contrast agents are dependent upon the local acoustic pressure. Thus, regional variations of acoustic pressure arising from myocardial anisotropy can result in nonuniform ''apparent perfusion'' in the absence of a perfusion defect. Measurements of regional variation of myocardial attenuation in the septum and lateral walls of excised sheep hearts are reported. Studies were carried out on hearts obtained immediately after slaughter from 8 healthy adult sheep. These measurements of attenuation were carried out in orientations analogous to those encountered in the apical four-chamber view, which is the preferred view for clinical studies of myocardial perfusion using contrast-agent-enhanced echocardiography. The video signal analysis method introduced by the University of Wisconsin group was used to analyze data obtained with a commercially available (ATL) clinical scanner using a linear array (L 7-4) operating in fundamental mode. Tissue mimicking phantoms were imaged under identical conditions. Normalizing sample data by reference data from the phantoms yields quantitative estimates of slope of attenuation. Results of these experimental measurements document significant transmural variations in attenuation that could lead to apparent perfusion anomalies in contrast-agent-mediated estimates of regional myocardial perfusion. [Work supported by NIHHL40302.]

10:30

3aBB12. Optimal transmit beam synthesis for coded-excitation pulse–echo imaging. Emad Ebbini and Mok-kun Jeong (Dept. of Elec. and Computer Eng., Univ. of Minnesota, 200 Union St., Minneapolis, MN 55455, emad@ece.umn.edu)

The use of code-fed arrays for pulse–echo ultrasonic imaging is gaining increased attention in medical applications. One of the most important applications is true real-time 3-D cardiac imaging where parallel processing of multiple-image lines is required to meet the stringent real-time constraints. The key to this approach is transmitting multiple codes to illuminate the region of interest (ROI), producing independent echo waveforms from different directions. We have examined several different scenarios in which the transmit aperture is divided into focused subapertures. Each subaperture is excited by one code and focused appropriately based on the definition of the ROI. For example, a 128-element array can be driven by eight independent codes on eight 16-element subapertures to illuminate a 6° ROI with f number of 2. The optimal solution is a sparse array pattern for each transmit subaperture (i.e., the subapertures overlap). Significant improvement in both lateral and range sidelobes of the system point-spread function (PSF) were achieved (compared to random transmit code selection). In this paper, we describe a new optimal synthesis of coded excitation transmit patterns. Examples of pulse–echo PSFs for the imaging system along with image reconstructions from a speckle-generating tissue-mimicking phantom will be given.

10:45

3aBB13. Simulation of tissue displacement using ultrasonic speckle tracking. Tjundewo Lawu and Mitsuhiro Ueda (Dept. of Intl. Development Eng., Tokyo Inst. of Technol., Tokyo, 152-8552 Japan)

The quantitative evaluation of ultrasonic displacement in tissue under externally applied forces is a necessary step in the field of ultrasonic diagnostics. In this study, the speckle tracking method is used to investigate the motion artifact produced by a rotating tissue. An analytic prediction of this motion artifact in relation to system characteristics (i.e., beam width, pulse duration, frequency, and phase of the signal) is presented. The tissue is modeled as a random array of numerous point scatterers, and rf signals are computed based on the convolutions between the transmission pulse and the impulse response of each scanning line. These analytical predictions are then compared to the computer simulation of the model. The results show that the artifact resulting from the speckle tracking method depends on the ultrasonic beam parameters. The method is quite general and can be extended to study the effects of other tissue motion, in particular, the deformation of tissue.

11:00

3aBB14. Statistical characterization of high-frequency ultrasonic scattering from human skin *in vivo*. Balasundar Raju and Mandayam Srinivasan (MIT, Rm. 36-788, 50 Vassar St., Cambridge, MA 02139, raju@mit.edu)

High-frequency ultrasound backscatter signals from forearm and fingertip skin *in vivo* of 20 normal human volunteers were analyzed in terms of their envelope statistics. Data obtained from several focal zones by axially translating a focused transducer were gated (100 ns) and Fourier transformed to compute magnitude of envelope at the center frequency of 30 MHz. It was found that the envelope histograms and SNR were significantly affected by the presence of occasional peaks that were unlikely to be from the background tissue. Hence a new parameter, the ratio of median to interquartile range (MQR) was used to evaluate its potential for tissue characterization. For all focal zone depths (up to 3 mm for forearm and 1.95 mm for fingertip), the MSR value was less than the Rayleigh limit of 1.30, indicating the presence of non-Rayleigh statistics. Significant differences in MSR were found between tissues close to the surface (0.3 mm) at the fingertip and forearm (0.67 ± 0.16 vs 0.97 ± 0.13 at 30 MHz, mean $\pm$ SD values are among subjects), which could correspond to anatomical differences at these locations, namely the increased thickness of the stratum corneum at the fingertip. These results indicate that non-Rayleigh statistics might aid in discriminating different skin tissues.

11:15

3aBB15. Detection of the malpositioned endotracheal tube via acoustic reflectometry: An *in vitro* study. David T. Raphael (Dept. of Anesthesiology, LAC-Univ. of Southern California Medical Ctr., Los Angeles, CA 90033 and UMDNJ-Robert Wood Johnson Med. School, NB, NJ)

Time-domain acoustic reflectometry can generate a one-dimensional area-distance profile of a cavity, such as the lung or esophagus. Can acoustic reflectometry be used to detect an endobronchial intubation? In an Institutional Review Board approved study, an endobronchial intubation in the right mainstem bronchus in an anesthetized patient was detected first by acoustic reflectometry, and the malposition confirmed by fiberoptic bronchoscopy. A study was initiated to determine the area-distance profile characteristics of an endobronchial intubation in an *in vitro* branching glass model. A symmetrical glass model (Witeg Scientific, Anaheim, CA), open at the proximal end, with two orders of bifurcation terminating in four closed distal branches, was studied. Area-distance (A-D) profiles were obtained with a customized computer-based acoustic reflectometer (Hood Labs., Pembroke, MA). With an endotracheal tube (ETT) attached to the reflectometer, area-distance profiles were obtained with the distal ETT in the following positions: the mid-''trachea,'' ''carina,'' ''mainstem bronchus,'' the second bifurcation, and the secondary ''bronchus.'' The area-distance profile features of an endobronchial intubation, as determined by this model study, are similar to those observed clinically in a patient.

11:30

3aBB16. Surface acoustic wave velocity in thin biotissue. Chiaki Miyasaka and Bernhard Tittmann (Penn State Univ., 212 Earth & Eng. Sci. Bldg., University Park, PA 16802)

We report a technique to characterize small portions on the order of a few microns of a biological tissue. The heart muscle was selected for specimen as an example of the soft material. The heart muscle was cut by a microtome and coated on a substrate. The thickness of the specimen was about 3 μm. Fused quartz was used as a substrate because its elastic properties are known and stable. A spherical lens was used to determine the position for the measurement. A cylindrical lens was used to measure velocities of the SAW waves within the specimen. 400 MHz was used for both imaging and measurement. The generation of the SAW waves was simulated by numerical calculations based on the wave propagation theory for layered media. As a result, the variation of the velocities representing the anisotropy of elastic properties was obtained and found to be about 7%.

11:45

3aBB17. Acoustic reflectometry for estimation of lung gas-phase volume during partial liquid ventilation in sheep. David T. Raphael (Dept. of Anesthesiol., LAC–Univ. of Southern California Medical Ctr., Los Angeles, CA 90033)

Acoustic reflectometry was studied as a possible tool to measure the lung gas-phase volume during partial liquid ventilation (PLV^{SM}). With IRB approval, healthy sheep (weight 10–22 kg) ($n=3$) were anesthetized, intubated, mechanically ventilated in the supine position, and underwent a low-lying tracheostomy. A base-line functional residual capacity (FRC) via a N_2 washout technique was measured. Perflubron (LiquiVent®, Alliance Pharmaceutical Corp., San Diego, CA) was instilled through the tracheostomy tube in successive aliquots, each consisting of 20% of measured FRC, up to 100% FRC. A Hood Labs acoustic reflectometer was used to obtain lung gas volume determinations for the FRC measurement and after perflubron instillations. The acoustic reflectometer was used to generate a cross-sectional area versus axial length curve. The subglottic lung gas-phase volume was calculated by integrating area over length. The reduction in the subglottic lung gas-phase volumes for each animal exhibited a strong linear correlation with the amount of perflubron instilled. Linear least-squares regression models run on the individual sheep data were all significant, with the largest p value being 0.0067. Acoustic reflectometry may be of use in the monitoring of lung gas-phase volume in PLV and respiratory disease.

12:00

3aBB18. Doppler ultrasound detection of shear waves remotely induced in tissue phantoms by focused ultrasound. Evgen A. Barannik, Sergii A. Girnyk, Volodymyr V. Tovstiak (Kharkiv Natl. Univ., Sq. 4 Svobody, Kharkiv, 61077 Ukraine), and Armen P. Sarvazyan (Artann Labs., North Brunswick, NJ 08902)

Remote generation of shear waves in tissues by radiation force of focused ultrasound is the basis of shear wave elasticity imaging (SWEI), a new acoustic method of medical diagnostics. Feasibility of SWEI was previously demonstrated in the experiments with optical and NMR detection of ultrasonically induced shear waves [Sarvazyan *et al.*, Ultrasound Med. Biol. **24**, 1419 (1998)]. In the present study the SWEI system with ultrasonic pulsed Doppler detection of shear waves was designed and tested using a range of gelatin-based phantoms and muscle tissue. Radiation force was generated by a focusing transducer of 8 cm diameter and 7 cm focal length. The carrier frequency was 1 MHz with intensity 1.85 W/cm2 at the surface of the transducer. The Doppler detection transducer operating at 3.5 MHz was installed in the center of the pumping transducer. The developed procedures of tissue motion detection and Doppler signal-processing algorithms based on the autocorrelation method allowed estimation of the velocity with accuracy better than 0.2 mm/s. The waveforms of the shear wave at various distances from the pumping ultrasonic beam axis were measured and the shear wave velocities were evaluated. The oscillation character of relaxation process for some phantoms and muscle tissue was shown experimentally.

WEDNESDAY MORNING, 6 DECEMBER 2000 CALIFORNIA SALONS 1 AND 2, 8:20 TO 11:40 A.M.

Session 3aEA

Engineering Acoustics: International Comparison of Calibration and Measurements

George S. K. Wong, Chair
Institute for National Measurement Standards, National Research Council, Montreal Road, Ottawa, Ontario K1A 0R6, Canada

Invited Papers

8:20

3aEA1. International calibration comparisons: Who benefits? George S. K. Wong (Inst. for Natl. Measurement Standards, Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

The mutual acceptance of acoustical calibrations and measurements between industrial countries is essential for international trade and the removal of no-tariff trade barriers. For example, the sound level or sound power emitted by a machine measured at the country of manufacture with certified instruments and methods in accordance with international standards, should be acceptable by the importing country without the requirement to duplicate the measurements. To achieve this mutual recognition, it is necessary for the exporting country to have proven capabilities via international comparisons and an unbroken chain of traceability from their national metrology institute to the machine shop level. Under the umbrella of the Bureau International does Poids et Mesures (BIPM), the Consultative Committee on Acoustics, Ultrasound and Vibration (CCAUV) has arranged international calibration comparisons, involving over 15 countries. The above comparisons require a lot of effort from each participating country. One may ask the question: Who is the beneficiary of international comparisons? The detailed answer is rather complex. In general, the results of International Calibration Comparisons provide confidence in the measurement capabilities of the participants. In the long term, the consumer is the ultimate beneficiary.

8:40

3aEA2. Key elements of mutual recognition arrangements. Klaus Brinkmann (Physikalisch-Technische Bundesanstalt Braunschweig, Bundesallee 100, 38116 Braunschweig, Germany, klaus.brinkmann@ptb.de)

Mutual recognition arrangements in the field of metrology are concluded between equal partners to promote the acceptance of each other's calibration certificates and test reports in order to facilitate global trade (''one-step-testing''). Acoustics is affected by this development both on the level of national metrology institutes, cooperating in organizations of the Metre Convention, and on the level of accreditation bodies and their accredited calibration and testing laboratories, cooperating in the International Laboratory Accreditation Cooperation (ILAC). Both arrangements are equivalent in the sense that they are based on mutual confidence between the partners, established mainly by the results of interlaboratory comparisons, uniform implementation of International Standards and peer assessments or accreditations.

9:00

3aEA3. Reference values for the sensitivity of standard accelerometers used in intercomparisons. David J. Evans, Stefan D. Leigh, and Beverly F. Payne (Natl. Inst. of Standards and Technol., 100 Bureau Dr., Stop 8221, Gaithersburg, MD 20899-8221, dje@nist.gov)

The National Metrology Institutes (NMIs) of five countries in North America and South America participated in an interlaboratory comparison involving the calibration of the magnitude of the sensitivity of three standard accelerometers. This comparison was performed by laboratories within the framework of the Interamerican Metrology System (SIM). One of the key values to be obtained in any interlaboratory comparison is an estimate of the reference values for the artifact being calibrated. Three statistical methods have been used to obtain candidate reference values and associated uncertainties from the SIM intercomparison data: an average of means method; a method based on the ISO *Guide to the Expression of Uncertainty in Measurement*; and a maximum likelihood method. Reference values and associated uncertainties obtained using the three methods are presented and compared, as well as useful graphical displays with resulting qualitative conclusions.

9:20

3aEA4. International key comparison of ultrasonic power measurements. K. Beissner (Physikalisch-Technische Bundesanstalt, 38116 Braunschweig, Germany, Klaus.Beissner@ptb.de)

One of the worldwide key comparisons performed at present under the auspices of the International Committee for Weights and Measures (CIPM) covers the time-averaged ultrasonic power emitted by an ultrasonic transducer into water. The ultrasonic power value is particularly important in characterizing the output of diagnostic and therapeutic medical equipment. Nine national metrology institutes participate in the comparison. The time schedule is from July 1999 to November 2001. The aim is to demonstrate the international consistency of calibration and measurement certificates. An ultrasonic standard transducer is circulated and power measurements are to be made at frequencies of 2, 6 and 10 MHz and at five power levels in the milliwatt and watt ranges. The measurement procedures applied and the main details of the technical protocol will be discussed.

9:40

3aEA5. International comparisons of hydrophone calibrations. Roy C. Preston, Stephen R. Robinson (Ctr. for Mech. and Acoust. Metrology, Natl. Physical Lab., Teddington, Middlesex TW11 0LW, UK, rcp@npl.co.uk), and Arnie L. Van Buren (Underwater Sound Ref. Div., Naval Undersea Warfare Ctr. Div., Newport, RI 02841)

In the global markets which are increasingly dominating industrial trade, international comparison of hydrophone calibrations is the primary mechanism of ensuring harmonization of underwater acoustics measurement standards across national borders. This paper describes the results of several recent comparison exercises initiated in Europe and the USA. These include both informal comparisons of hydrophone calibrations of a bilateral nature and formal comparisons involving a number of countries organized under the auspices of regional or international metrology organizations. Completed projects which will be described include a Euromet/European Commission project involving a total of 12 participants from seven European countries and a comparison between the Russian Federation and China. In these cases, results will be given and the important conclusions summarized. Finally, a relatively new comparison project will be described which has been set up under the auspices of the Consultative Committee for Acoustics, Ultrasound and Vibration (CCAUV) of the Bureau International des Poids et Mesures (BIPM) in Paris. This key comparison has now been agreed to for the kilohertz frequency range 1–500 kHz and will involve the circulation of three hydrophones. Countries participating in the comparison include the UK, China, Germany, Russian Federation, and USA, with Canada expected to join soon.

10:00–10:15 Break

10:15

3aEA6. Importance of the free-field calibration of microphones. Victor Nedzelnitsky (Natl. Inst. of Standards and Technol. [NIST], Sound Bldg. [233], Rm. A147, 100 Bureau Dr., Stop 8221, Gaithersburg, MD 20899-8221, Victor.Nedzelnitsky@nist.gov)

Numerous regulatory requirements, standards, and product characterization and quality control procedures important for industry, commerce, health, and safety rely on practical measuring instruments including sound level meters, personal sound exposure meters (noise dosimeters), and standardized measuring microphones. Such instruments are usually large enough that the effects of diffraction must be considered throughout a significant portion of the operating frequency range. These effects cause differences between the free-field, the pressure, and the diffuse-field sensitivities of instruments, particularly at high frequencies. Differences can also result from other phenomena, such as the effects of a microphone's static (barometric) pressure equalization vent, especially at low frequencies. Different calibration procedures have evolved, so that an appropriate calibration is available for each of the field types:

free, pressure, and diffuse. Critically important free-field procedures and some of the most accurate and widely used diffuse-field procedures should be, but are not always, traceable to primary free-field calibrations of microphones by the reciprocity method. Future international key comparisons at the highest level of accuracy should include determinations of the free-field sensitivities of laboratory standard microphones such as IEC type LS2 by the reciprocity method.

10:35

3aEA7. Reference sound source calibration at various temperatures and site altitudes. Angelo J. Campanella (Campanella Assoc. & Acculab, 3201 Ridgewood Dr., Columbus, OH 43026)

Adjustments are made for atmospheric temperature and barometric pressure when calibrating a sound power reference sound source (RSS) and when applying an RSS in the substitution method to evaluate an unknown sound power. Experience with Hemianechoic calibration of a dozen RSS units outdoors at various temperatures near sea level and a special test series on one RSS unit at three higher sites up to 9927 feet and applying a pistonphone reference to verify microphone and sound-level meter stability are described. According to these varied temperature results and the theoretical work of others [Allan Powell, J. Acoust. Soc. Am. **32**, 982–990 (1960), Eqs. (41)–(42), after Lighthill], it appears that the mechanical fan-type RSS is a multipole—chiefly quadrupole and octupole—type sound source. It was apparent, according to these varied altitude results, that emitted sound pressure and sound power are both linear functions of barometric pressure. Experimental results will be presented. Practical applications to determine actual sound power, to report normalized sound power, and to predict actual sound pressure and sound power expected at a remote site under nonstandard conditions are described.

Contributed Papers

10:55

3aEA8. Model-based method for improvements in the calibration of ultrasonic power measurements at low-power levels. Lixue Wu (Acoust. Standards, Inst. for Natl. Measurement Standards, Natl. Res. Council Canada, Ottawa, ON K1A OR6, Canada)

The measurement uncertainty in the calibration of ultrasonic power measurements at low-power levels was reduced by a model-based signal-processing method. The ultrasonic power measurement system is based on the radiation force balance method according to IEC 61161. Outputs of an electronic balance and the excitation voltage were simultaneously sampled and recorded. A computer-controlled coaxial switch that utilizes data lines of the parallel port of a desktop computer was used for power-on and power-off control. Power-on and power-off intervals were finely controlled by the desktop computer. The Electroacoustic radiation conductance, defined as the ratio of the ultrasonic power of a transducer and the square of the rf voltage at its input, was calculated from the recorded data. During data processing, a simple model for drifting and switching was proposed and a least-mean-squares algorithm was used to estimate parameters of the model. Good measurement reproducibility was also observed that resulted in the reduction of total measurement uncertainty.

11:10

3aEA9. Boundary-element simulation used in the design of a sound intensity calibrator. Johan Gramtorp, Vicente Cutanda, Anders Eriksen, and Erling Sandermann Olsen (Bruel & Kjaer, Skodsborgvej 307, 2850 Naerum, Denmark)

The recent introduction of hand-held intensity measurement equipment has created a growing need for verification of the measurement equipment in the field. The key feature for easy field use is the possibility of calibrating the two-microphone intensity probe without dismounting the spacer. Working models were built, but none of them worked to more than approximately 3 kHz. Compliance with IEC 61043 requires the calibrator to work up to 7.1 kHz. A lot of modifications on these models were tried, but none of them worked. The development project was close to being stopped. A boundary-element model of the sound intensity calibrator was built. It verified the measured result from the working models. Based on the boundary-element model the best type and position of the sound source and the optimum dimensions of the calibrator cavity were found. Measurements on a final model verified the simulation results. The calibrations could now be made without dismounting the spacer.

11:25

3aEA10. Calibration of electret microphone cartridges for application in a multichannel acoustic array. Allan J. Zuckerwar, William M. Humphreys, and Scott M. Bartram (NASA Langley Res. Ctr., M.S. 493, Hampton, VA 23681)

Electret condenser microphone cartridges, having the advantages of small size, low cost, few lead wires, low output impedance, and acoustic specifications comparable to those of air condenser microphones, are well suited for application in multichannel acoustic arrays. Amplitude and phase calibration of electret cartridges by two methods is described. In the first, a test cartridge, reference microphone, and driver are inserted into a coupler, whereby the calibration takes place by a comparison method. In the second method, the calibration is performed in a commercial acoustic calibrator (B&K type 4226), which was modified by the manufacturer to permit the phase calibration. Further, a special adapter was fabricated to seat the cartridge at the proper position in the calibrator coupler. Amplitude and phase calibrations of 80 electret microphones were performed. The calibrated microphone cartridges were flush-mounted in an 80-channel array to measure dynamic surface pressures. The array was installed in a backward-facing step model and tested in a low-speed wind tunnel. An associated data acquisition system performed conditioning and digitization of the signals from each of the microphones. Preliminary results using the array to characterize the pressure distribution in the separated flow region downstream of the step will be reported.

Session 3aMU

Musical Acoustics: Asian Musical Instruments and Traditions I

James P. Cottingham, Chair
Department of Physics, Coe College, Cedar Rapids, Iowa 52402

Chair's Introduction—8:25

Invited Papers

8:30

3aMU1. The acoustics of the Asian free reed mouth organs. James P. Cottingham (Phys. Dept., Coe College, Cedar Rapids, IA 52402, jcotting@coe.edu)

Mouth-blown instruments using a free reed coupled to a pipe resonator have a long history in China, Japan, and throughout Southeast Asia. The sheng, khaen, and bawu have been studied experimentally and theoretically as typical representatives of this family of instruments. Acoustical measurements made include studies of reed vibration and impedance measurements of the pipes. Particular attention has been paid to the coupling of the reed vibration with the pipe resonator. The sheng employs a free reed at one end of a closed tube with a conical-cylindrical cross section. The khaen employs an open tube of effective length L, with the reed located at approximately $L/4$. The bawu is a closed cylindrical pipe with the free reed at one end, in which the effective acoustical length is varied by the use of tone holes. The playing frequency of each pipe of the sheng or the khaen is typically slightly above both the resonant frequency of the pipe and the natural frequency of the reed. In the bawu, on the other hand, both the pipe resonance and sounding frequency are normally well above the natural reed frequency, resulting in a striking difference in tone quality.

9:00

3aMU2. Some acoustics of the shakuhachi—and of the shakuhachi player's face. Joe Wolfe and John Smith (School of Phys., Univ. of New South Wales, Sydney NSW 2052, Australia, J.Wolfe@unsw.edu.au)

The shakuhachi is an end-blown Japanese flute. Like other flutes, it is open at the place of excitation, so it operates at minima of the acoustic impedance. The jet interacts with both the bore impedance and the radiation impedance, which is baffled by the player's face. Large changes in the relative geometry of the instrument and the face contribute to the flexibility in pitch and timbre that are important elements of the traditional playing style, and which gives the instrument much of the expressiveness for which it is renowned. We report measurements of the impedance spectra $Z(f)$ of the bore, and of the baffled radiation load imposed on the jet. $Z(f)$ of the instrument differs from that of the Western flute family because the shakuhachi has neither the narrow chimney of the Western flute, nor the short resonator upstream of the jet. These influence the overall form of $Z(f)$ for the instrument. Other features of $Z(f)$ are explained by the tone hole position and geometry. The $Z(f)$ of the radiation field varies for different playing positions, which has important effects on the tuning of the minima and on the spectral response.

9:30

3aMU3. The Klais organ in Kyoto: A comparison of the acoustics of the Japanese organ stops to the traditional Japanese instruments. Jonas Braasch (Institut für Kommunikationsakustik, Ruhr-Universität, Bochum, Germany) and Christian Ahrens (Ruhr-Universität, Bochum, Germany)

The Klais organ in Kyoto is unique among organs of the world in having a section with Japanese organ stops. Although common organ stops are sometimes named *shakuhachi* in Japan, the Klais organ is the only one that has, besides a French and a German section, a Japanese section containing four specially designed stops imitating the Japanese instruments: *shakuhachi*, *shinobue*, *sho*, and *hichiriki*. While the *shakuhachi* and *shinobue* stops use flue pipes, free-reed pipes are used for the *sho* and the *hichiriki*. Sound recordings of single tones were made of these Japanese stops and analyzed regarding the sound spectra and the attack transients, as well as the variation of the fundamental frequency during the attack transient. For comparison, sound recordings of traditional instruments were made at the Geijutsu Deigaku University of Tokyo. The results show that the sound characteristics of the traditional instruments could be imitated quite well. The free-reed pipes work well not only to imitate the free-reed instrument *sho*, but also to imitate the double-reed *hichiriki*. This organ is a good example in which the long-neglected free-reed stops are more suitable than the commonly used beating reed stops. [Work supported by DFG.]

10:00

3aMU4. Chinese string instruments 400 B.C.–200 A.D. Bo Lawergren (Hunter College of the City Univ. of New York, 695 Park Ave., New York, NY 10021, bo.lawergren@hunter.cuny.edu)

When the tomb of Marquis Yi of Zeng (deceased 422 B.C.) was opened in 1978, a large number of musical instruments came to light. Many of those have since been well studied—in particular the bronze and stone chimes—but the string instruments (various types of zithers) have not. This year the author published the first Western language study of the zithers from Zeng and contemporary Chinese sites (B. Lawergren, ''Strings,'' in *Music in the Age of Confucius*, edited by J. F. So, Washington, DC), and it forms the basis

of this survey. Some types of Zeng zithers were precursors of instruments (*qin* and *zheng*) that became important during the last two millennia of Chinese music. Other types (*se* and the *zhu*) died out after some centuries. The most interesting precursor is that of the classical *qin*. The latter has been held in highest esteem since at least 200 A.D., imbued with deep philosophical meaning, and considered an indigenous invention *par excellence*. But several features of the precursor appear to be foreign, probably from central or western Asia. A reproduction will be used to demonstrate sounds, tuning systems, and concepts of the Zeng *qin*.

10:30

3aMU5. The Burmese harp: Its scale and tunings. Robert M. Williamson[a] (Emeritus, Dept. of Phys., Oakland Univ., Rochester, MI 48309)

The 16 strings of the Burmese horizontal, arched harp are tuned in pentatonic octaves over a range of two-and-a-half octaves. A brief description of the harp, its history, and examples of its classical music will precede a discussion of the results of pitch measurements of harp tones recorded in Mandalay in the mid-20th century of the late Muriel C. Williamson [*The Burmese Harp* (Southeast Asia Press, DeKalb, IL, 2000)]. Prior to about 1960, the seven-tone Burmese harp scale could be accurately described by the notes CDE*(E) F G A *(A) B*, where the starred notes were about a quarter tone (50 cents) lower than the corresponding Western notes. The pitches of the notes E* and A* could be varied when they occurred as stopped tones of the harp. The distinctive Burmese notes became Western within a few decades. The remarkably precise tunings of master harpists sometimes showed octaves stretched by about 15 cents. Previous measurements of Burmese fixed-pitch instruments show them to have had somewhat tempered version of the harp scale. [a] Present address: 1189 Pond Cypress Drive, Virginia Beach, VA 23455.

11:00

3aMU6. How are two tones produced simultaneously in throat-singing? Seiji Adachi (ATR Info. Sci. Dept., 2-2 Hikaridai Seika, Kyoto, 619-0288 Japan) and Masashi Yamada (Osaka Univ. of Arts, Higashiyama Kanan, Osaka, 585-8555 Japan)

The most remarkable characteristic of throat-singing is that one singer simultaneously produces two distinct tones—drone and melody. The drone has a low and sustained pitch, and is sung with a pressed voice. The melody has a whistlelike tone, whose pitches are much higher than the drone. It was conjectured that the production of the melody was caused by the resonance of the vocal tract, but has never been proved. This study clarifies that the high melody pitch is produced by the pipe resonance of the rear cavity in the vocal tract. This is derived from acoustic investigations on a throat-singer's vocal tract measured by magnetic resonance imaging. Four different shapes of the vocal tract are examined, with which the melody pitches of F6, G6, A6, and C7 are sung along with the F3 drone of a specific pressed voice. The second formant frequency calculated from each tract shape is close to the melody pitch within an error of 36 cents. Sounds are synthesized by convolving a glottal source waveform with transfer functions calculated from the vocal tract shapes. Two pitches are found to be successfully perceived when the synthesized sounds are listened to.

WEDNESDAY MORNING, 6 DECEMBER 2000 CALIFORNIA SALON 4, 7:55 TO 11:45 A.M.

Session 3aPA

Physical Acoustics: Thermoacoustics I

Gregory W. Swift, Chair
Los Alamos National Laboratory, K764, Los Alamos, New Mexico 87545

Chair's Introduction—7:55

Contributed Papers

8:00

3aPA1. Frequency response of a nonlinear acoustical resonator of arbitrary shape. Mark F. Hamilton, Yurii A. Ilinskii, and Evgenia A. Zabolotskaya (Dept. of Mech. Eng., Univ. of Texas, Austin, TX 78712-1063, zhenia@mail.utexas.edu)

The frequency response of a nonlinear acoustical resonator is investigated analytically and numerically. First, natural frequencies of a resonator whose shape is close to cylindrical are studied analytically by perturbation. An asymptotic formula is derived for the natural frequencies as a function of resonator shape. The solution at first order in the perturbation shows that each natural frequency can be shifted independently via appropriate spatial modulation of the resonator wall. Numerical calculations for resonators of different shapes establish the limits of the asymptotic formula. Second, the nonlinear interactions of modes in a nearly cylindrical resonator are investigated with Hamiltonian formalism as a basis for studying wave interaction in a resonator of arbitrary shape. A nonlinear frequency response function for the fundamental mode is derived by assuming that nonlinear interaction occurs only with the second harmonic. Bent tuning curves are predicted which lean up or down depending on whether the second-harmonic frequency is above or below the corresponding natural frequency, respectively. A fully nonlinear 1-D numerical code is used to verify the analytical result for the frequency response. [Work supported by ONR.]

8:15

3aPA2. Two-dimensional instability study of a thermoacoustic prime-mover. Ivan Delbende (Univ. of Paris 6, LIMSI, BP 133, 91403 Orsay Cedex, France)

A numerical model aimed at investigating the development of instability in a thermoacoustic prime-mover is presented. It is based on a discretization of the two-dimensional compressible Navier–Stokes equations, and does not rest on the plane pressure wave assumption. The computational domain is made of a single stack interplate space extended up to both resonator ends. The working gas is assumed to be initially at rest, and

3a WED. AM

a longitudinal temperature gradient is enforced. The linear instability properties of the system are then given by the eigenmodes of the discretized evolution operator linearized in the vicinity of the above basic state. Temporal growth rate and oscillation frequency are thus obtained for all thermoacoustic modes, and the prime-mover instability onset is determined as a function of the various parameters (mainly temperature gradient magnitude and mean pressure). As far as *linear* instability is concerned, this formulation is attractive, since it avoids any direct numerical simulation and any problem involved by the presence of two very different time scales.

8:30

3aPA3. Experimental instability study in a thermoacoustic prime-mover. Emmanuel Bretagne, Ivan Delbende, and Maurice-Xavier François (Univ. of Paris 6, LIMSI, BP 133, 91403 Orsay Cedex, France)

An experimental linear stability study of a thermoacoustic prime-mover is performed for different values of the mean pressure between 0.5 and 10 bars. The damping rate is carefully obtained as a function of the temperature gradient $|\nabla T|$ enforced along the stack, up to the instability onset at $|\nabla T|_c$. These results are then confronted with the predictions of a numerical model based on Rotts' theory used with complex frequencies, in which the prime-mover is seen as a feedback loop in the electrical analogy. The experimental results are found to reasonably comply with Rott's theory as soon as the mean pressure exceeds 2 bars. Below this value substantial discrepancies are found, as confirmed by the work of Yazaki. Above the instability onset, the saturated wave amplitude is measured as a function of $|\nabla T|-|\nabla T|_c$ for fixed pressure. The bifurcation to the nonlinear saturated wave has been tentatively determined as subcritical, although thermal inertia effects make it look supercritical.

8:45

3aPA4. Fluctuation before onset of oscillations in a thermoacoustic device. Young Sang Kwon, Orest G. Symko, and Karin Durrant (Phys. Dept., Univ. of Utah, 115 S. 1400 E., Rm. 201, Salt Lake City, UT 84112-0830)

The onset of oscillations in a resonant device where a temperature gradient produces a sound is a fundamentally interesting problem as it consists of a transition where random gas motion, when biased, produces an almost pure sound. This was studied in small 1/4-wave resonant tubes containing a fibrous stack across which a temperature gradient was maintained by means of heat exchangers at each end. The temperature gradient biases the random motion of the gas to trigger a standing wave in the resonator. The acoustic spectrum was measured as the temperature gradient approached the onset for oscillations. Factors affecting the transition to oscillation consist of stack gain, quality factor of the resonator, magnitude of the temperature gradient, and acoustic load. The directed diffusion of the gas along the stack leads to a series of sharp pressure pulses whose stochastic behavior triggers resonant oscillations in the tube. Onset temperature differences less than 50 °C have been observed and this could further be reduced by a suitable choice of parameters affecting the onset. This is an example of a thermal ratchet engine biased by a temperature gradient.

9:00–9:15 Break

9:15

3aPA5. A graphical software application for design and simulation in thermoacoustic research (DSTAR). Thomas J. Hofler (Dept. of Phys., Naval Postgrad. School, Monterey, CA 93943, tjhofler@nps.navy.mil), Eric Purdy, and Scott Curtis (U.S. Navy)

A graphical application called DSTAR has been written for the 32-bit Windows environment, for the purpose of thermoacoustic heat engine analysis. Written in C++ and MFC, the application uses pull-down menus and tabbed dialog boxes, and has additional graphical features such as automatic plotting of thermoacoustic state variables and geometry output in CAD file format. DSTAR supports a variety of variable units including dimensionless units, which are extremely convenient in the design phase of an engine. Current versions of the software are freely available at http://cooler.physics.nps.navy.mil/hofler/.

9:30

3aPA6. Development of a miniature thermoacoustic refrigerator. Thomas J. Hofler and Jay A. Adeff (Dept. of Phys., Naval Postgrad. School, Code PH/HF, Monterey, CA 93943, tjhofler@nps.navy.mil)

A miniature thermoacoustic refrigerator is being developed for the purpose of cooling integrated circuits below their failure temperature in hot environments. Work has been done on a piezoelectric acoustic driver operating at 4 kHz. A simple refrigerator has been built and tested that uses one atmosphere of air as a working medium, is less than 3 in long, and has produced 12 °C of cooling. A more advanced pressurized refrigerator is under development. [Work supported by DARPA and Rockwell Science Center.]

9:45

3aPA7. Redesign of a small thermoacoustic refrigerator. Tamra S. Underwood, Bryan S. Morley, Benjamin L. Kuritz, and Ralph T. Muehleisen (Dept. of Civil, Environ., and Architectural Eng., Univ. of Colorado, Boulder, CO 80309)

At the Fall 1999 meeting of the Acoustical Society of America, the design and initial performance of a small, electrically powered thermoacoustic refrigerator was presented. The refrigerator has nominal dimensions of 5 in. long and 1 in. diameter with a target of 10 W of cooling over a 25 °C temperature span using 2-bar helium as the working gas. Because of limitations in the acoustic transducer, performance was limited. A new design with a more powerful acoustic transducer is presented. The performance of the actual device will be compared to the predicted values. [Work supported by the Office of Naval Research.]

10:00

3aPA8. A large solar/heat-driven thermoacoustic cooler. Reh-lin Chen and Steven L. Garrett (Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804, rxc132@psu.edu)

Based on the success of an earlier solar-powered thermoacoustics prime mover which used a direct-illumination stack and no hot-side heat exchanger [Chen and Garrett, Proc. 16th Int. Cong. Acoust., Vol. II, 813–814 (1998)], a large solar/heat-driven thermoacoustic cooler was designed and fabricated. Target cooling powers of 10 to 60 W, over a 25-deg temperature span, were based on a thermal input power of 150 to 600 W. To concentrate the required amount of solar power on an 11-cm-diameter ceramic stack, a 10-ft diameter fiberglass parabolic dish, used for satellite TV, has been converted by gluing aluminized Mylar™ on its surface over a 2-m diameter. A two-axis coordinated solar tracking system, driven by two computer-controlled motors, has produced the required 600 W of solar power to illuminate the hot side of the stack for a maximum of 3 h. Measured performance of the solar refrigerator will be compared to DELTAE models. [Work supported by the Office of Naval Research.]

10:15

3aPA9. Performance improvements for the solar powered, thermoacoustically driven, thermoacoustic refrigerator. Jay A. Adeff and Thomas J. Hofler (Naval Postgrad. School, Phys. Dept. Code PH, Monterey, CA 93943)

The world's first completely solar powered, thermoacoustically driven, thermoacoustic refrigerator has been improved with increased cooling power and higher efficiency. A larger 24-in.-diam Fresnel lens provides increased heat power to focus sunlight at 550 °C onto the hot end of a 1-in. reticulated vitreous carbon prime mover stack. The high-intensity

sound waves produced by the prime mover are used to power a thermoacoustic refrigerator. The cooling power has been substantially improved and the temperature span has been increased from 18 to 30 °C with the cold end at 0 °C.

10:30–10:45 Break

10:45

3aPA10. A two-microphone thermoacoustic impedance tube. Timothy Simmons, Richard Raspet, and Robert Hiller (Dept. of Phys. and Astron., and Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677, tgsimmon@olemiss.edu)

An experimental arrangement has been developed to explore the thermoacoustic properties of stack materials. The apparatus may be viewed either as a thermoacoustic engine with two microphones mounted in the tube wall, or as a two-microphone impedance tube that accommodates thermoacoustic elements. A sample stack material is placed between two heat exchangers and specific acoustic impedance measurements are taken as a function of frequency and at low-drive amplitude (to avoid nonlinear effects). The design is such that nearly all physical elements and operational parameters can be varied fairly easily. The ''tube'' is composed of three (flanged) pipe sections of different length, which independently contain a loudspeaker, stack and heat exchangers, and a termination. The test is performed with and without a temperature gradient imposed across the stack. The tube may be pressurized—allowing for different gases held at various ambient pressures—while the relative positioning of the elements within the tube may be varied. Hence a very wide range of testing conditions may be achieved. [Work sponsored by ONR.]

11:00

3aPA11. Thermoviscous functions of wire mesh and RVC stacks. Ralph T. Muehleisen and C. Walter Beamer (Dept. of Civil, Environ., and Architectural Eng., Univ. of Colorado, Boulder, CO 80309)

Thermoacoustic stacks made of stacked wire-mesh elements or reticulated vitreous carbon (RVC) are becoming popular because of their performance, cost, and ease of construction. The thermoviscous Rott functions f_ν and f_κ [related to the porous media function $F(\lambda)$] have been measured for wire mesh and RVC stacks. The thermoviscous functions were determined from measurements of the characteristic impedance and propagating wave number of the stacks. The measurements were then used to develop empirical models for the stacks. The measurement technique, measurement results, and emprical models are presented. [Work supported by the Office of Naval Research.]

11:15

3aPA12. High-amplitude viscous effects in single pores. Andi Petculescu and Larry Wilen (Dept. of Phys. and Astron., Ohio Univ., Clippinger 251, Athens, OH 45701, apetcule@helios.phy.ohiou.edu)

Recently, we reported on a new technique to measure viscous effects of single pores [A. Petculescu and L. Wilen, J. Acoust. Soc. Am. **107**, 2819 (2000)]. The method involves a simple lumped-element analysis of a compliant region in parallel with the pore. The pore impedance is determined by subtracting off the measured compliance. With two lengths of the same pore and a second subtraction, one can determine the impedance of the uniform middle part of the pore as well as the impedance of the pore ends. We have extended our measurements to higher amplitudes to look at nonlinearities in the pore. We will discuss how we can determine that the nonlinear effects are principally due to the pore ends, and how we measure these effects. Specifically, results will be presented for the amplitude dependence of the impedance as well for as the generation of higher harmonics. Different end geometries will be considered (such as rounded versus sharp tube ends) and the implication for jet pumps will be addressed. [Work supported by ONR.]

11:30

3aPA13. Thermoacoustic oscillation around a plate in a nonlinear standing wave field. Huang Dongtao, Guo Qing, and Zhu Zhichi (Dept. of Eng. Mech., Tsinghua Univ., Beijing 100084, PROC, dongtaohuang@263.net)

A stack of plates is a most basic and important component of thermoacoustic engines and thermoacoustic refrigerators. Analysis of thermoacoustic oscillation around a plate is a basis of analysis of the mechanism of these machines, but only the works in a linear standing wave field were published. On the other hand, it is shown from some experiments that the thermoacoustic efficiencies of these machines will be increased in the nonlinear standing wave field, but any quantitative result has not been found yet. In this paper a numerical simulation of thermoacoustic oscillation around a plate in a standing wave tube where the sound-pressure level of the sources is from 90 to 170 dB has been presented in order to discuss the nonlinear effects quantitatively. Two conclusions have been obtained: (1) thermoacoustic efficiencies are increased with the increase of the source slowly below 150 dB, moderately from 150 to 160 dB, and rapidly above 160 dB. (2) There exists a rule of multiplying growth of high-order harmonics of oscillation temperature level, which is very similar to the rule of multiplying growth of high-order harmonics of oscillation sound-pressure level in a nonlinear standing wave tube.

Session 3aSA

Structural Acoustics and Vibration: Vibrations of Structures

Courtney B. Burroughs, Chair
Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16802

Contributed Papers

9:00

3aSA1. A unified method for free vibrations of elastically restrained plates arbitrarily loaded with springs and masses. W. L. Li and M. Daniels (United Technologies Carrier Corp., A+R Bldg., Carrier Pkwy., Syracuse, NY 13221)

A simple and unified Fourier series method is presented for the vibration analysis of a class of plates that are simply supported along a pair of opposite edges and elastically restrained along the others. The plates can also be arbitrarily loaded with springs and masses. The transverse displacements of the plates are expressed as the superposition of a Fourier sine series and a simple cubic polynomial. The polynomial is introduced to take care of all the discontinuities with the original displacement and its relevant derivatives along the edges so that the Fourier series now simply represents a residual (or conditioned) displacement which is continuous and has at least three continuous derivatives. The excellent accuracy and convergence of the current method are demonstrated through various numerical examples.

9:15

3aSA2. Weak and strong couplings between a master oscillator and a set of satellite oscillators. G. Maidanik (Carderock Div., Naval Surface Warfare Ctr., Code 7030, 9500 MacArthur Blvd., West Bethesda, MD 20817-5700)

The distribution of resonance frequencies of the satellite oscillators *in situ* is imposed. This imposition demands that the stiffness control term in the satellite oscillator's impedance and the stiffness control term in the coupling, of this oscillator to the master oscillator, be supplemental. Moreover, the distribution is aligned in ascending order; namely, $\omega r \leqslant \omega q$; $q=(r+1)$; $1 \leqslant r \leqslant (R-1)$, where (ωr) and (ωq) are the resonance frequencies of the (r)th and (q)th satellite oscillators, respectively, and (R) is their number. In addition, the numbers of resonance frequencies of the satellite oscillators, on either side of the resonance frequency (ωo) of the master oscillator, are rendered equal. The couplings are assigned mass, stiffness, and gyroscopic control terms. Examples of the influence that the satellite oscillators collectively have on the response of the master oscillator, under various coupling forms and strengths, will be cited.

9:30

3aSA3. On the use of Ritz series as an alternative to field equations for modal analysis of continuous systems. Jerry H. Ginsberg (G. W. Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405)

Mathematical analysis of a continuous system's transient response commonly relies on modal properties obtained as an eigensolution of the field equations. Such solutions become cumbersome if the system has attached springs, masses, or substructures, in which case the eigenvalue problem requires mathematically connecting subdomains of the continuum with continuity conditions and Robin boundary conditions. Derivation of the modal orthogonality conditions, knowledge of which is required for implementing the expansion theorem, is quite intricate in such cases. The Rayleigh–Ritz method is a widely used alternative for identifying modes. Representing the displacement field with a Ritz series in conjunction with Hamilton's principle, which is commonly known as the method of assumed modes, offers a generalization of the Rayleigh–Ritz method that can also be employed to analyze forced response. The present paper will demonstrate that if one appropriately alters their viewpoint, and performs a few minor intermediate steps, it is possible to use the Ritz series formalism to obtain the same results (mode functions, orthogonality properties, and the expansion theorem) as those obtained by solving field equations. Such a derivation solely requires knowledge of the kinetic and potential energy functionals, and therefore is much easier to implement for multiply connected systems.

9:45

3aSA4. Natural frequencies of nonaxisymmetric vibration of prolate spheroidal shells. Sabih I. Hayek (Dept. of Eng. Sci. and Mech., Penn State Univ., University Park, PA 16802) and Jeffrey E. Boisvert (Naval Undersea Warfare Ctr. Div. Newport, Newport, RI 02841)

The equations of motion for nonaxisymmetric vibration of prolate spheroidal shells of constant thickness were derived using Hamilton's principle. The thin shell theory used in this derivation includes three displacements and two changes of curvature. The effects of membrane, bending, shear deformations, and rotatory inertias are included in this theory. The resulting five partial differential equations are self-adjoint and positive definite. The nonaxisymmetric modal solutions are expanded in a doubly infinite series of comparison functions. These include associated Legendre functions in terms of the prolate spheroidal angular coordinate, and circular functions of the circumferential coordinate. The natural frequencies and the mode shapes were obtained by the Galerkin method for each circumferential mode. Numerical results were obtained for several shell thickness-to-length ratios ranging from 0.005 to 0.1, and for various diameter-to-length ratios, including the limiting case of a spherical shell. [Work supported by Office of Naval Research and the Navy/ASEE Summer Faculty Program.]

10:00–10:15 Break

10:15

3aSA5. Predicted and measured structural intensities in the lexan and aluminum INCE standard ribbed panels. Stephen Hambric (ARL/Penn State, P.O. Box 30, State College, PA 16804), Richard Szwerc, and Henry Chang (NSWCCD, West Bethesda, MD 20817-5700)

Comparisons of predicted and measured structural intensities in vibrating lexan and aluminum ribbed panels are made. Structural intensities are predicted using finite-element models composed of plates and beams and the McPOW (mechanical power) software. Wave decomposition methods were used to estimate bending wave intensities by postprocessing velocity fields on the lexan and aluminum ribbed panels measured using a scanning laser Doppler vibrometer (SLDV). Each panel was driven at one rib with a point drive and suspended with soft springs to simulate free boundary conditions. Intensity fields are compared for each panel's (7,2) mode. The intensity fields are normalized by the power input to the panels by the point drive. The predicted intensity fields for both panels, when normalized by the power input, are nearly identical. The measured intensity fields

are similar to each other and to the predicted fields, but show some differences. The discrepancies are due to measurement uncertainties and manufacturing differences. The results help mutually validate the prediction and measurement approaches.

10:30

3aSA6. On the use of the Sturm sequence to evaluate modal density. Philip J. Shorter (Vibro-Acoust. Science, Inc., 12555 High Bluff Dr., Ste. 310, San Diego, CA 92130, pj.shorter@vasci.com)

The modal density of a structural-acoustic subsystem is usually obtained analytically by considering the dispersion of various propagating wave types. Closed form expressions are available for the modal densities of simple beams, plates and shells (with curvature in one or two directions). However, one is often interested in subsystems with complex geometry which may possess inhomogeneous material and physical properties. The classical asymptotic formulations are not appropriate for such subsystems and numerical methods are often adopted. The most straightforward approach is to perform a finite-element-based modal analysis and count the number of eigenvalues that fall within various frequency bands. However, the computational expense associated with solving the full eigenproblem is often prohibitive. Significant computational savings can be made by employing the Sturm sequence property to evaluate the modal density. This paper describes the approach in more detail and provides a numerical example.

10:45

3aSA7. Test stand for measuring the vibration of impact-type hand-held tools. Douglas Reynolds and Jeff Markle (Ctr. for Mech. and Environ. Systems Technol., Univ. of Nevada, Las Vegas, 4505 Maryland Pkwy., Las Vegas, NV 89154)

Vibration from hand-held impact tools is a major cause of hand-arm vibration syndrome (HAVS) in industrial workers. Evaluating the effectiveness of engineering methods for reducing tool vibration in impact tools has been hampered because of the inability to measure the impact vibration of the working end of the tools. A test fixture has been developed that can be used to measure the impact vibration at the working end of a hand-held impact tool. These measured values can be used to determine whether or not the claimed vibration reduction of a tool is related to an effective engineering modification to the tool without reducing the productivity of the tool. Measurements were made on a rivit hammer to determine the effectiveness of a vibration-reducing mechanism applied to the tool attachment section at the working end of the tool. This mechanism reduced the vibration at the handle of the tool by 80%.

11:00

3aSA8. Cable strum self-noise cancellation for sonar towed arrays. Vincent E. Premus (MIT Lincoln Lab., 244 Wood St., Lexington, MA 02420)

Nonacoustic self-noise observed on marine seismic streamers and towed sonar arrays represents a serious problem for acoustic source detection at low frequency. Towed array self-noise, also known as cable strum, consists of mechanical vibrations induced by vortex shedding. Transverse vibrations in the array body subject each hydrophone pressure head to local accelerations. The resultant acoustic response can be several orders of magnitude stronger than the water-borne acoustic signals of interest. In this paper, a beamspace, time-domain adaptive signal processing architecture for the coherent rejection of broadband nonacoustic self-noise is presented. The approach is based on the recognition that most vibrational modes of a towed array propagate at phase speeds substantially less than those of acoustic signals in the water column. This property supports the formation of a signal-free, strum reference using the same sensor that samples the acoustic data. The approach removes the need for additional measurement channels, such as accelerometers or strain gauges, to independently sense the undesirable distortions introduced by cable strumming. The phenomenology underlying flow-induced self-noise for towed arrays is discussed and characterized using k-w analysis. An overview of the method is presented and cancellation performance is demonstrated using snapshots of passive sonar towed array data. [Work sponsored by the Department of the Navy, under Air Force Contract F19628-95-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the U.S. Air Force.]

Session 3aSC

Speech Communication: Alvin M. Liberman and the Development of Scientists

Doug H. Whalen, Chair
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511-6695

Invited Papers

9:00

3aSC1. Alvin M. Liberman and the development of scientists: Chairman's introduction. D. H. Whalen (Haskins Labs., 270 Crown St., New Haven, CT 06511)

Alvin M. Liberman contributed to the theoretical field of speech for half a century, and was still contributing at the time of his death in January of 2000. The papers in this session will celebrate some of those contributions, but will also highlight another aspect of his career that made him a powerful force: He brought out the best scientific effort of all those around him—his students, his colleagues, his children, and (even perhaps especially) his critics. Liberman's influence will continue for the lifetimes of those who knew him, and no doubt longer through his impressive body of work.

9:05

3aSC2. Alvin Liberman, infant research, and the speech module. Peter D. Eimas (Dept. of Cognit. and Linguistic Sci., Brown Univ., Providence, RI 02912) and Joanne L. Miller (Northeastern Univ., Boston, MA 02115)

In 1957, Alvin Liberman offered a highly controversial statement based on the experimental work at Haskins Laboratories, known as ''the motor theory of speech perception.'' Over time, the theory evolved to presume that the mechanisms underlying speech processing were biologically determined and specialized for the perception and production of speech. Speech perception was in effect not an auditory process, at least beyond the initial reception of speech, but rather part of a species-specific processing system that made language possible. We consider this view in light of selected aspects of our research on infant speech perception, which was driven in large part by the work of Liberman and his associates. We believe that the evidence in its entirety supports Liberman's idea of a species-specific processor or module that permits the acquisition and use of language, including much of the relatively early processing of speech.

9:20

3aSC3. Alvin M. Liberman and the development of Midwestern speech scientists. James J. Jenkins (Dept. of Psych., Univ. of South Florida, Tampa, FL 33620-8200)

In the summer of 1968, Alvin Liberman came to the University of Minnesota to teach a special seminar in speech perception for the Center for Research in Human Learning. His efforts with the gifted students of the Center not only produced a piece of original research on the perception of speech, but also initiated a productive relationship between the University and Haskins Laboratories. Faculty and students from three different academic departments trekked back to New Haven to make stimuli and profit from the intensely interacting pattern of collaborative work at Haskins Laboratories. This largess was possible because of the special relationship between Haskins and NICHD that supported our use of the Haskins facilities. This relationship was typical of Al's outreach to interested scientists everywhere. At the University of Minnesota alone, over the next decade the relationship produced numerous masters theses and doctoral dissertations. More importantly, it facilitated the training and development of almost a dozen speech scientists who are continuing to advance the field of research to which Al devoted his professional life.

9:35

3aSC4. Functional role of efferent feedback to the inner ear. M. Charles Liberman (Eaton-Peabody Lab., Mass. Eye and Ear Infirmary and Dept. of Otology and Laryngology, Harvard Med. School, Boston, MA 02115)

With Alvin Liberman as a father, a career in scientific research was the only one I seriously considered. However, I did stray as far as the study of hearing rather than speech. One of the major interests of my laboratory is the functional role of the neuronal system of efferent feedback to the inner ear, known as the olivocochlear pathway. One part of this olivocochlear pathway projects to the outer hair cells and constitutes a sound-evoked reflex which, when activated, reduces the sensitivity of the cochlea and raises thresholds to sound. In a series of neurophysiological studies in cat and guinea pig, we have shown that, paradoxically, this efferent feedback can act to improve the detection of transient signals in steady background noise and that the presence of an intact efferent pathway also protects the ear from acoustic injury due to overly intense sound.

9:50

3aSC5. Categorical perception and the convergence of social learning. Mark Y. Liberman (Dept. of Linguist., Univ. of Pennsylvania, 619 Williams Hall, 36th and Spruce St., Philadelphia, PA 19104-6305)

About 50 years ago, Alvin Liberman and his colleagues at Haskins Laboratories discovered ''categorical perception'' of certain phonetic dimensions, where discrimination is hardly better than identification. They proposed this as one piece of converging evidence for a special perceptual mode for speech, part of a species-specific evolutionary adaptation for articulate language. This special sensory-motor module, they argued, makes human speech so uniquely efficient at transmitting discrete symbol sequences in sound. Lively arguments about all aspects of this reasoning continue to this day. This paper broadens the discussion beyond perception in individuals, by suggesting a role for categorical perception in enabling communities of speakers to form and maintain consensus about the pronunciation of tens of thousands of morphemes. Through computer simulations, simple and plausible assumptions will be shown under which the ''pronouncing dictionaries'' of the members of a speech community will converge rapidly from random starting points. Among these assumptions, a form of categorical perception plays a key role. With it, simulations converge to a consensus from which deviations occur rarely and hardly ever spread. Without it, individual pronunciation beliefs wander chaotically across time in the phonetic space, and at a given time, differences in belief increase rapidly with social distance. The role of partly categorical perception, as in so-called ''magnet effects,'' will also be addressed.

10:05–10:15 Break

10:15

3aSC6. A biological basis for writing and reading. Michael Studdert-Kennedy (Haskins Labs., 270 Crown St., New Haven, CT 96511)

As Alvin Liberman taught us many years ago, speech is a complex acoustic code on the phonological units of its message, but writing is a simple cipher on that message. Yet we learn to speak more readily than we learn to read and write—a fact central to Liberman's biological view of speed. Must there not, however, also be a biology of writing and reading? Why is writing possible at all? What property enables language, alone among systems of animal communication, to be transduced into an alternative, but no less efficient, perceptuomotor modality? The outline of an answer emerges when we recall that writing systems, like speech itself, represent not meaning, but the intrinsically meaningless units, whether syllables or phonemes, on which the hierarchy of language is raised. Such units lend themselves to transduction not only because they have no meaning, but also because, as vehicles of communicative ''parity'' between speakers and listeners, they are neither sensory nor motor but cognitive and abstract. The nature and origin of these particulate entities at the base of language was the topic of a paper that Liberman and I were planning at the time of his death. The paper was to have been called ''A Biological Basis for Writing and Reading.''

10:30

3aSC7. The importance of parity; its implications for understanding speech perception. Carol A. Fowler (Haskins Labs., 270 Crown St., New Haven, CT 06511, University of Connecticut, and Yale Univ.)

Early in his career, Alvin Liberman obtained experimental findings that he describes in his book, *Speech: A Special Code*, as ''an epiphany.'' The findings revealed that listeners track speakers' articulations. From these findings and subsequent others, Liberman developed his motor theory of speech perception. In addition, however, he set out to understand why listeners track articulation. His explanation was that humans evolved a phonetic system jointly responsible for producing coarticulated speech and for recovering intended phonetic segments from coarticulated signals. Underlying the claim that production and perception are linked in this way is the idea of parity. In Liberman's interpretation, parity is the requirement that, for language to serve as a major component of human communication systems, listeners and talkers must agree on what set of perceivable human actions can count as components of a linguistic message. In addition, and more concretely, in their conversational interactions, listeners and talkers must typically achieve a relation of parity between messages sent and received. This paper will focus on how this parity requirement rationalizes the body of findings that listeners track articulatory gestures.

10:45

3aSC8. Articulatory gestures in the motor theory. Ignatius G. Mattingly (Haskins Labs., 270 Crown St., New Haven, CT 06511 and Dept. of Linguist., Univ. of Connecticut, Storrs, CT 06269)

According to Alvin Liberman's motor theory, the objects of speech perception are ''articulatory gestures.'' But, what is an articulatory gesture? In early statements of the theory, this phrase seems to mean simply the movements of a single articulator, as idealized in traditional impressionistic-phonetic terms such as ''tongue raising.'' But, the notion of the articulatory gesture was greatly enriched over the years as Liberman and his colleagues took account of evidence from experiments in both speech perception and speech production. In the most recent statements of the motor theory, the articulatory gesture has become a temporary modification of the geometry or the excitation of the vocal tract, typically involving more than one articulator, often global rather than local. To Liberman and his colleagues, the human ability to distinguish and identify various such gestures from the available acoustical and optical data, despite extensive temporal overlap, appeared more remarkable than ever.

11:00

3aSC9. Perceiving speech: A developmental perspective. Janet F. Werker (Dept. of Psych., Univ. of British Columbia, Vancouver, BC V6T 1Z4, Canada)

The question of how we so instantaneously and effortlessly perceive the rapidly changing and highly encoded speech signal was one of the guiding questions in Alvin Liberman's research career. From the original presentation of the ''motor theory'' through the rich body of research it generated and its subsequent revisions, he posed for the field a challenge to view speech in a radically different way from what he called the ''dominant'' view. On my first trip to Haskins Laboratories as a young graduate student, Alvin Liberman listened carefully to and took seriously the (naive) research questions I was attempting to ask about young infants, and challenged me to consider whether my approach was sufficiently rich to capture the essence of speech perception. That challenge continues to guide my work. In this talk I will present research that I (and others) have conducted with infants and young children in an attempt to understand the nature of speech perception in the young infant, and how that changes as a function of experience—both heard and seen—with the native language.

11:15

3aSC10. Alvin M. Liberman's legacy: Haskins Laboratories. D. H. Whalen (Haskins Labs., 270 Crown St., New Haven, CT 06511)

Alvin M. Liberman came to Haskins Laboratories in June 1944 to help create a reading machine for the blind. Under Franklin S. Cooper, his group did create the first such system but, more importantly, Liberman led the way in creating an environment where the basic issues of speech and reading could be addressed by an assemblage of specialists which could not reasonably be supported by any one academic institution. The technological tools that were first available only at Haskins—controlled speech synthesis on the Pattern Playback, EMG recording, and dichotic presentation in particular—were the primary attractors, but the collaborations continued long after the equipment became less than unique. Haskins embodied the scientific ideal of the free exchange of information, across disciplinary boundaries, to a greater extent than most other institutions. A large force behind it all was Liberman's insistence on the highest technical expertise backing up theoretically interesting experiments. He was rightfully proud of the fact that Haskins results always replicated, whether the interpretation was accepted or not. The atmosphere of scientific rigor continues at Haskins, and will remain so as long as we can continue to uphold the high standards that Liberman set for us.

WEDNESDAY MORNING, 6 DECEMBER 2000 BRIGANTINE ROOM, 10:25 TO 11:45 A.M.

Session 3aSP

Signal Processing in Acoustics: Acoustical Imaging

David H. Chambers, Chair
Lawrence Livermore National Laboratory, P.O. Box 808, L-154, Livermore, California 94551

Chair's Introduction—10:25

Contributed Papers

10:30

3aSP1. Inversion methods for ultrasonic travel-time tomography. James G. Berryman (Lawrence Livermore Lab, P.O. Box 808 L-200, Livermore, CA 94551)

Ultrasound travel-time tomography in a laboratory setting, where data can be collected without limited views in 2-D and/or 3-D imaging, has many advantages over seismic travel-time tomography in the earth. When samples to be imaged have relatively high contrast, some new methods that might be called ''velocity spectroscopy'' and ''smart backprojection'' become viable for fast and accurate imaging of any number of inclusions, as long as they are not too small and not too dense in the imaging plane. For low contrast, some simple methods have also been developed for ray bending analyses that are relatively quick and more accurate than straight ray methods, such as filtered backprojection. Applications of these methods to synthetic and real data will be discussed.

10:45

3aSP2. Recent developments in 3-D SAFT reconstruction applied to laser-ultrasonics. Daniel Levesque, Alain Blouin, Christian Neron, and Jean-Pierre Monchalin (Industrial Mater. Inst., Natl. Res. Council of Canada, 75 de Mortagne Blvd., Boucherville, QC J4B 6Y4, Canada, daniel.levesque@nrc.ca)

The Synthetic Aperture Focusing Technique (SAFT) has been used to improve both the detectability and the lateral resolution of conventional ultrasonics. This method has recently been applied to laser-ultrasonics, an ultrasonic technique based on the generation and detection of ultrasound with lasers. Actual 3-D SAFT reconstruction is still considered computationally intensive and the pseudo-3-D SAFT, stacking line SAFT images together, is more commonly used. A recent development in 3-D SAFT reconstruction is toward improvements of frequency domain SAFT. In addition to increasing speed, the improved method (F-SAFT) includes temporal deconvolution to enhance both axial and lateral resolutions, control of the aperture to improve signal-to-noise ratio, as well as spatial interpolation in subsurface planes. These last two features also allow a reduction in the sampling requirements to further decrease both data ac-

quisition time and processing time, therefore making the technique more attractive for industrial use. All the above operations are well adapted to the frequency domain calculations and embedded in the F-SAFT processing. The performance of F-SAFT reconstruction will be illustrated using laser-ultrasonic data taken from test samples with flat-bottom holes.

11:00

3aSP3. Acoustic coherence imaging through the atmosphere. Jonathan W. Benson (SAIC / Demaco, 100 Trade Ctr. Dr., Ste. 303, Champaign, IL 61820, jonathan.w.benson@saic.com)

The process of making images of acoustic scenes through the atmosphere has been investigated. Representations of the acoustic power with angle and the power spectrum with angle can be formed. The performance of interferometric imaging methods based on the Van Cittert–Zernike theorem has been evaluated using data collected in a series of field experiments. An array of microphones received the far-field emissions from a pair of large loudspeakers that reproduced broadband random noise signals. Both the loudspeakers and the array are located near the ground. All imaging is done in the horizontal plane. Pairs of microphones in the array are correlated to form measurements of the complex coherence. An image is then formed via an inverse Fourier transform operation. The field data show that making accurate acoustic images outdoors is difficult due to the effects of turbulence. The image distortion, signal-to-noise ratio, and variance associated with the coherence measurements were calculated as a function of the average wind speed. Even at moderate wind speeds (6 m/s) turbulence can cause an order of magnitude increase in the variance of coherence measurements. Guidelines were drawn that define the expected performance based upon the observed wind speed. Methods for image enhancement are discussed.

11:15

3aSP4. Optimal source distribution for virtual acoustic imaging. Takashi Takeuchi and Philip A. Nelson (ISVR, Univ. of Southampton, Highfield, Southampton SO17 1BJ, UK, tt@isvr.soton.ac.uk)

When binaural sound signals are presented with loudspeakers, the system inversion involved gives rise to a number of problems such as loss of dynamic range and a lack of robustness to small errors in control performance. Regularization, often used to design practical filters, also results in poor control performance around ill-conditioned frequencies. These problems for such systems are investigated and this has resulted in the proposal of a new system concept. The system overcomes these fundamental problems by means of a conceptual pair of monopole transducers whose span varies continuously as a function of frequency. The underlying theoretical principle is described in detail. The significance is that all of the above problems that are associated with the multi-channel system inversion are solved by using this principle. The limitations with this principle are also made clear in terms of the operational frequency range. Several examples of practical solutions that can realize a variable transducer span by discretization are also described. The discretization expands the operational frequency region to be used with only a little decrease in performance. This principle is extremely useful and practical because a single transducer which can cover the whole audible frequency range is not currently available.

11:30

3aSP5. Ultrasonic fault detection by processing of signals from fixed transceiver system. Khan Mohammad Mahmud (Div. of Appl. Phys., Grad. School of Eng., Hokkaido Univ., Sapporo, 060-8628 Japan, mah@eng.hokudai.ac.jp) and Ryoji Ohba (Hokkaido Univ., Sapporo, 060-8628 Japan)

A nonscanning method for an ultrasonic flaw detection technique is described. The system employs an M-sequence modulated ultrasonic wave as the excitation signal. Irrespective of the number of faults to be detected, the data acquisition system predeterminedly consists of a transmitter and a few receivers to be fixed at convenient locations. The cross-correlation function (CCF) between the original M-sequence and the demodulated received sequence indicates the presence of faults in the form of sharp peaks. In order to clearly detect the expected peaks corresponding to the faults in the CCF, new signal processing techniques are proposed. Experimental results (on metal plate) demonstrate the capability of the system and confirm the feasibility of the approach under heavy ambient noise condition. An algorithm, only on the basis of travel time of the signal, to determine the exact location of a fault is described. Synchronous moving average and coincidence multiplication process are found to be very effective in eliminating the false peaks from the CCF while increasing the SNR significantly. The former can be easily modified for the purpose of next generation online signal processing of an ultrasonic NDT signal. [The authors would like to express their gratitude to the Ministry of Education, Japan, for support with the scholarship as well as the funds needed for this research.]

Session 3aUW

Underwater Acoustics: Modeling

Chris T. Tindle, Chair
Applied Research Laboratory, Pennsylvania State University, Box 30, University Park, Pennsylvania 16804-0030

Chair's Introduction—7:55

Contributed Papers

8:00

3aUW1. On the use of the sonar equation approximation when computing the response of a target in shallow water. Angie Sarkissian and Louis R. Dragonette (Naval Res. Lab., Washington, DC 20375-5350, angie@aqua.nrl.navy.mil)

The sonar equation may be used to compute the scattering response of a target in shallow water by approximating the target to be a point scatterer. The approximation significantly simplifies the interaction of the scatterer with the medium. In the more general case, since the field produced by a source arrives at the scatterer through multipaths, it is incident on the target at various directions. The response of the scatterer must be computed at all of the incident angles to obtain the correct scattered field in shallow water. Scattering results computed using the sonar equation are compared to the more correct solution for a ribbed cylindrical shell placed in shallow water for various monostatic and bistatic geometries to examine the validity of the use of the sonar equation at various frequencies and geometries. [Work supported by ONR.]

8:15

3aUW2. Sonar performance in turbid and bubbly environments. Simon D. Richards (Underwater Sensors & Oceanogr. Dept., Defence Evaluation and Res. Agency, DERA Winfrith, Dorchester, Dorset DT2 8XJ, UK) and Timothy G. Leighton (Univ. of Southampton Highfield, Southampton SO17 1BJ, UK)

The presence of both solid particles and gas bubbles in coastal waters may have a significant effect on sound propagation, particularly at high frequencies, and may therefore be partially responsible for the observed variability in high-frequency sonar performance in shallow waters. Suspended particles increase volume attenuation through the processes of thermo-viscous absorption and scattering. Microbubbles also attenuate sound through viscous and thermal dissipation and scattering. The presence of microbubbles in the water column may also modify the sound speed and result in a dispersive medium. The effects of dilute suspensions of solid particles on the sound speed may generally be neglected for practical sonar applications. Algorithms for estimating the acoustic absorption coefficient and the speed of sound in water containing suspensions of fine mineral particles and populations of microbubbles have been developed. These have been incorporated into a sonar performance prediction model based on the ray method. Results are presented from this model which show that, for populations of suspended mineral particles and microbubbles which are commonly encountered in shallow-water environments, the predicted sonar performance is significantly different from that predicted when the additional effects are not taken into account. []

8:30

3aUW3. High-frequency propagation modeling with the parabolic equation. David M. Fromm, Michael D. Collins (Naval Res. Lab., Washington, DC 20375), and Guy V. Norton (Stennis Space Center, MS 39529)

It is generally believed that the parabolic equation method rapidly becomes impractical above a few hundred Hertz for ocean acoustics problems. Results will be presented to illustrate that the split-step Padé algorithm [J. Acoust. Soc. Am. **100**, 178–182 (1996)] is practical out to ranges of tens of kilometers for frequencies up to at least on the order of 10 kHz. The presentation will include a discussion of modifications to the parabolic equation model to handle effects that are usually neglected at lower frequencies, such as rough surfaces, internal waves, and volume attenuation in the water column. The examples will involve propagation in surface ducts and interaction with the seafloor. [Work supported by the Office of Naval Research.]

8:45

3aUW4. A wide-angle parabolic equation for advected acoustic waves. Joseph F. Lingevitch, Michael D. Collins, Dalcio K. Dacol (Naval Res. Lab., Washington, DC 20375), Joel C. W. Rogers (Polytechnic Univ., Brooklyn, NY 11201), and William L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180)

Parabolic equations are derived by factoring operators in elliptic equations. For acoustics problems involving an ambient flow, the elliptic equation contains two depth operators that do not commute with each other. Due to this difficulty in factoring the operator, the parabolic equations that have been derived for this problem are limited to small grazing angles (i.e., energy that propagates nearly horizontally). A wide-angle parabolic equation for advected waves will be described and examples will be presented to demonstrate its accuracy. As in previous investigations of this problem, the Mach number is assumed to be small. The derivation is based on the introduction of a new dependent variable that is related to the acoustic pressure by an amplitude operator that can be implemented using either of two approximations. A zeroth-order approximation (in Mach number) is useful in some cases since the error is local in range. Greater accuracy can be achieved using a first-order approximation, which can be expressed in terms of the derivative of an operator with respect to a parameter and implemented using a difference formula. [Work supported by the Office of Naval Research.]

9:00

3aUW5. A new stable formulation of an elastic parabolic equation. Wayne Jerzak, William L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180-3590), Michael D. Collins, and Joseph F. Lingevitch (Naval Res. Lab., Washington, DC 30375-5330)

The parabolic equation method can be used to model acoustic-wave propagation in elastic media. Current implementations do not accurately match the shear stress between two elastic layers, for which the expression involves a second-order depth derivative. This inaccuracy could be large

in range-dependent problems. A reformulation of the elastic equations in terms of new variables permits all interface conditions to be handled accurately since no expressions contain second- or higher-order depth derivatives. However, the implementation in the new variables is found to be unstable because some evanescent modes are allowed to grow. Higher-order evanescent modes can be suppressed using a rational approximation to the depth operator with a polynomial in the denominator having a higher degree than the numerator. The lower-order evanescent modes can be controlled by imposing constraints on the choice of coefficients in the approximation. Examples are presented to show the accuracy and stability of the new parabolic equation in shallow-water environments. [Work supported by ONR.]

9:15

3aUW6. Computation of the acoustical wave-fronts in a phase space. Nick E. Maltsev (MVM Intl., 10678D Maplewood Rd., Cupertino, CA 95014, Maltsev_Nick@msn.com)

The computation of acoustical wave-fronts in 2-D and 3-D coordinate space can be performed by the computation of rays and approximation of the fronts between the rays by some interpolation formula. The amount of neccessary rays increases with distance and rapidly exhausts memory and performance of the computer. It happens due to the extremely sophisticated and unsmooth nature of the ray picture. Due to the fact that trajectories of a dynamic system do not intersect in phase space, appropriate rays and fronts surface in phase space is smooth and can be easily approximated by a smaller number of rays, and this number increases very slow with distance. This number can be reduced by involving equations in variations, which generate tangent vectors to this surface. Formally, dimension of phase space is a doubled dimension of coordinate space. Using some properties of ray equations, the dimension of the phase space in the 2-D coordinate case can be reduced from 4-D to 3-D and in 3-D coordinate case from 6-D to 5-D. This approach opens new ways of fast global and local estimates of sound field and for inverse problems in underwater acoustics. This report produces a set of numerical and analytical examples.

9:30

3aUW7. Wave-front modeling. Chris T. Tindle (Appl. Res. Labs., Penn State Univ., Box 30, University Park, PA 16804-0030)

Wave-front modeling offers a direct method of finding pulse waveforms in both deep and shallow water. The wave-front is found by tracing rays to a fixed range and finding the ray depth as a function of starting angle. Interpolation on the wave-front provides the amplitude and arrival time at the receiver depth. Phase is determined by the ray history. The expressions for the phase and amplitude are derived directly from the wave equation using WKB approximations for solutions to the depth-separated equation. Caustics are handled by treating rays as pairs and involve Airy functions which describe the smooth transition between insonified and shadow regions. On the shadow side of a caustic the rays have complex angles and travel times. Ray-tracing for complex angles can be avoided and waveforms in the shadow can be found approximately by fitting the phase function near the caustic.

9:45–10:00 Break

10:00

3aUW8. Evaluation of schemes for interpolating eigenray properties between updates in acoustic propagation modeling. Joseph A. Clark (CDNSWC, Code 734, Bethesda, MD 20084, clarkja@nswccd.navy.mil)

Acoustic simulators can model the propagation of sound in the ocean by computing a set of eigenrays between sources and receivers. A matrix of features characterizing each eigenray is computed using environmental data. The eigenray matrices are then used to modify the source signals received by each receiver. Temporal variations, introduced into the eigenrays by motions of the sources and receivers, require that the eigenray matrices be updated. Some form of interpolation between updates is needed in order to avoid the introduction of artifacts into the simulated signals. Doppler effects must also be appropriately reproduced in the data. This talk describes a method for evaluating interpolation schemes by performing transient analyses of the simulated signals. Graphical displays used to examine the data streams will be shown. Visual and aural indications of artifacts produced by poor interpolation schemes and difficult modeling conditions will be identified. Evaluations of some improved interpolation schemes will be presented.

10:15

3aUW9. A locally conformal method for modeling rigid boundaries in the FDTD method. John B. Schneider and Julius G. Tolan (School of Elec. Eng. and Computer Sci., Washington State Univ., Pullman, WA 99164-2752)

The finite-difference time-domain (FDTD) method is a numerical technique that can be used to solve a wide variety of problems via time-domain simulations. A low-cost method has previously been presented for accurately modeling continuously varying pressure-release surfaces [Schneider *et al.*, J. Acoust. Soc. Am. **104**, 3219–3226 (1998)]. The overall simplicity of the FDTD method was preserved by requiring modifications only in the vicinity of the pressure-release boundary. Thus, the implementation of such things as sources or absorbing boundary conditions can be left unchanged. This paper describes a technique, which is similar in spirit but different in detail, that permits accurate modeling of continuously varying rigid boundaries. The technique is simple to implement, has low computational cost, and provides substantial improvements over a traditional ''staircased'' representation of the boundary. Accuracy is established by comparing results to those for canonical scatterers. [Work supported by ONR.]

10:30

3aUW10. The use of complex images in acousto-elastic propagation problems. John A. Fawcett (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada)

In many propagation and scattering problems it is necessary to compute a half-space (for example, water over a seabed) Green's function for a variety of source/receiver positions. For example, in the boundary integral equation method (BIEM), the kernel of the integral equation involves the Green's function, and numerically the Green's function must be evaluated at a large number of source/receiver positions. The distances involved may range from near-field, where the Green's function will have a singular behavior, to the far-field. In this paper a method from the electromagnetics literature [e.g., Vitebskiy *et al.*, IEEE Trans. Antennas Propag. **44**, 143–151 (1996)] is described which reduces the usual wave number integral representation to a small number of image sources with complex positions. This representation is valid over a large range of source/receiver distances and hence is appropriate for efficiently evaluating the Green's function for many source/receiver positions. Numerical examples are presented illustrating the accuracy of the method for both fluid and elastic seabeds.

10:45

3aUW11. Monte Carlo model for underwater ambient noise fields. Daniel L. Hutt, Andrew L. Rosenfeld, and Paul C. Hines (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 2Z7, Canada, daniel.hutt@drea.dnd.ca)

Three-dimensional ambient noise fields possessing realistic cross-spectral properties have been simulated using the Monte Carlo method. In the model, weighted monochromatic sources are distributed about an array of sensors to create the desired directional characteristics of the noise field. The hydrophones can be positioned arbitrarily to form any desired array structure and orientation. Noise signals are represented in the frequency domain where phase shifts due to propagation are most naturally applied. Examples of simulated cross spectra for a variety of noise fields will be presented. Good agreement with theory is found for cases in which analytical solutions are possible.

11:00

3aUW12. Simulation of underwater intensity vector measurement: Effect of errors and noise. Daniel L. Hutt, Paul C. Hines, and Victor W. Young (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada, daniel.hutt@drea.dnd.ca)

Pressure and pressure-gradient signals from an array of hydrophones can be processed to yield the three-dimensional sound-intensity vector of an acoustic field. In underwater applications, the technique may prove useful for near-field measurements of acoustic intensity and characterization of ambient noise fields. Unfortunately, the accuracy of the magnitude and direction of the calculated intensity vector is degraded by many factors including: fluctuations in the received signal, the presence of ambient noise, system noise, phase error, and errors in channel sensitivity. To investigate the effect of these errors, a numerical simulation of the pressure field due to signal sources and ambient noise was developed. The model is used to calculate the variance of intensity measurements under different noise conditions and to simulate the effects of channel gain and phase imbalance. The results are found to be in good agreement with theoretical expressions. The simulation was also used to compare the performance of a six-element orthogonal intensity array and a four-element tetrahedral intensity array in terms of gain against noise and variance in estimated direction to a sound source.

11:15

3aUW13. Numerical results for an approximate form of the nonlocal small slope approximation scattering strength. Shira L. Broschat (School of Elec. Eng. and Computer Sci., Washington State Univ., P.O. Box 642752, Pullman, WA 99164-2752)

Voronovich introduced the nonlocal small slope approximation (NLSSA) as a generalization of the small slope approximation to explicitly include nonlocal interactions. He showed that the NLSSA generally accounts for double scattering in the high-frequency limit. Broschat and Thorsos presented numerical results for the lowest-order NLSSA scattering strength for two-dimensional pressure-release surfaces. Their results agreed well with Monte Carlo integral equation results and, in particular, were better than the higher-order SSA results at low forward grazing angles. However, the computational cost was extremely high, and results were unobtainable at low grazing angles. In this paper, we discuss the results obtained by making an *ad hoc* approximation to the lowest-order NLSSA scattering cross section that reduces the computational complexity of the integration substantially. In addition, we discuss the difficulties of the numerics and how they are handled. Numerical results for the scattering strength are presented for 2-D pressure-release surfaces. For the cases presented, the NLSSA and its approximate form give virtually the same results. Also, results at low grazing angles are obtainable and accurate with the approximate form. [Work supported by ONR.]

11:30

3aUW14. Spectral integral representations of multistatic scattering from sediment volume inhomogeneities. Henrik Schmidt (Dept. of Ocean Eng., MIT, 77 Massachusetts Ave., Cambridge, MA 02139) and Kevin D. LePage (SACLANT Undersea Res. Ctr., La Spezia, Italy)

A perturbation approximation for multistatic scattering from sediment volume inhomogeneities has been developed using a Fourier–Bessel spectral integral representation of the source and receiver Green's functions. This approach is particularly efficient for modeling scattering from horizontally isotropic, three-dimensional distributions of scatterers in waveguides with arbitrary horizontal layering, and is accurate for scatterers in fast bottoms near the critical angle and for scatterers in layers with background sound-speed gradients. The theory has been implemented in the OASES code and the model has been exercised over broad bandwidths in order to explore the temporal and angular evolution of bistatic scattering for a variety of environmental and experimental scenarios. Examples show that the scattered field received on a bistatic vertical line array shows a distinct evolution in time and angle which can be simply related to the source–receiver geometry and the location of the contributing scatterers at any given time. Similarly, the horizontal angular evolution of bistatic scattering shows a time-angle trajectory which may be easily interpreted using similar geometrical arguments. [Work partially supported by ONR.]

WEDNESDAY MORNING, 6 DECEMBER 2000 BIG CANYON I/II ROOM, 9:00 A.M. TO 12:00 NOON

Meeting of Accredited Standards Committee (ASC) S2 on Mechanical Vibration and Shock

to be held jointly with the

U.S. Technical Advisory Group (TAG) Meeting for ISO/TC 108 Mechanical Vibration and Shock (and Subcommittees ISO/TC 108/SC1, SC2, SC3, SC5, and SC6)

R. J. Peppin, Chair S2
5012 Macon Road, Rockville, Maryland 20852

D. J. Evans, Vice Chair S2 and Chair of the U.S. Technical Advisory Group (TAG) for ISO/TC 108, Mechanical Vibration and Shock
National Institute of Standards and Technology (NIST), 100 Bureau Drive, Stop 8221, Gaithersburg, Maryland 20899-8221

Accredited Standards Committee S2 on Mechanical Vibration and Shock. Working group chairs will present reports of their recent progress on writing and processing various shock and vibration standards. There will be a report on the interface of S2 activities with those of ISO/TC 108 (the Technical Advisory Group for ISO/TC 108 consists of members of S2, S3, and other persons not necessarily members of those Committees), including plans for future meetings of ISO/TC 108 and/or its Subcommittees.

Scope of S2: Standards, specifications, methods of measurement and test, and terminology in the fields of mechanical vibration and shock, and condition monitoring and diagnostics of machines, but excluding those aspects which pertain to biological safety, tolerance and comfort.

Session 3pAA

Architectural Acoustics: Amphitheater Acoustic Design and Sound Control for Nearby Communities

Dana S. Hougland, Cochair
ACODA Acoustics, LLC, 9603 Orchard Drive, Englewood, Colorado 80111-3503

Leslie D. Blomberg, Cochair
Noise Pollution Clearinghouse, P.O. Box 1137, Montpelier, Vermont 05602-1137

Chair's Introduction—1:00

Invited Papers

1:05

3pAA1. Amphitheater noise, a community perspective. Leslie D. Blomberg (Noise Pollution Clearinghouse, Box 1137, Montpelier, VT 05601, les@nonoise.org)

Amphitheaters and amplified music often provoke strong community reaction. Leaders of approximately 20 community organizations near existing or proposed amphitheaters were surveyed, and three were selected for case studies. Their experience provides insight into the affected public's perception of the concert industry, acoustical consultants, ordinances, and regulations.

1:25

3pAA2. Outdoor concert sounds—Searching for acceptability criteria for communities adjacent to amphitheaters. William J. Cavanaugh (Cavanaugh Tocci Assoc., Inc., 327 F Boston Post Rd., Sudbury, MA 01776)

Most municipal, state, and federal codes and regulations are inadequate for protecting residential communities adjacent to concert amphitheaters from annoyance due to sound of high-level amplified music, as well as from other concert-related sounds. From numerous consultations involving siting of new and evaluating existing amphitheaters, the author has evolved a criterion which appears to predict with sufficient accuracy the likely response of residential occupants to amphitheater sounds. This experience suggests that if the intrusions represented by the 1-percentile A-weighted levels are not more than 5 dB above the pre-existing (i.e., nonconcert) background represented by the 90-percentile levels, the intrusions are sufficiently immersed in and masked by the normal background and most communities are satisfied. However, some neighbors find the mere audibility or detectability of amphitheater sounds objectionable and measurement is at best difficult where the sound levels are of the same order of magnitude as the normal community background sound. The experience in developing and applying this signal-to-noise based criterion to several amphitheater projects, one of which has been in operation for over 15 years, is reviewed.

1:45

3pAA3. Evolution of metrics used in ''real world'' concert sound level management systems. Richard G. Cann (Grozier Technical Systems, Inc., 157 Salisbury Rd., Brookline, MA 02145, rcann@grozier.com)

With the ever increasing popularity and proliferation of outdoor amphitheaters and the advent of more powerful amplification systems serving high-energy performing groups, nearby residents are inevitably threatened by higher sound levels. To address these concerns, acoustical consultants have specified monitoring systems that must use simple metrics for the management and control of music facility sound emissions. These metrics must be pragmatic; the visiting artist and sound engineer, the concert promoter, the concert facility management personnel, and even local regulatory personnel must be able to easily understand and interpret their meaning. Furthermore, the management system must accommodate rapid real-time implementation of necessary management control measures and store accurate, permanent, electronic records for future reference. This paper describes the evolution of the Grozier sound level management systems that use simple, accurate metrics developed for several amphitheater venues with diverse local conditions and requirements.

2:05–2:15 Break

2:15

3pAA4. Using time-variant electronic architecture to improve outdoor acoustics. Steve Barbar (LARES Assoc., 9170 Rte. 108, Ste. 404, Columbia, MD 21045)

The outdoor environment offers little of the acoustical support provided by the traditional architecture and surface treatments found in good concert halls or opera houses. While these treatments can be used as part of a shell or enclosure where the musicians perform, incorporating architectural treatments that improve listening enjoyment for the audience, or return energy for the musicians, is much more difficult. Sound reinforcement systems used for classical music or opera outdoors can assist in improving the impact of the direct sound when used carefully. Such systems, however, can do little to improve the sense of envelopment and warmth that are

an important part of the listening experience for this type of performance. Recent developments in technology have produced a system that uses time variance to reduce the effects of acoustic feedback. These systems have a significant advantage in gain before coloration over previous methods (typically 12 dB or more). When used outdoors, this system provides reflected and reverberant energy for both musicians and the audience. Additionally, this energy can help to mask noise intrusion from sounds outside the designated seating area. The following is what we have determined to be important parameters for the successful implementation of such a system.

2:35

3pAA5. Community noise reduction through sound system renovation, a case study. David E. Marsh (Pelton Marsh Kinsella, 1420 W. Mockingbird Ln., Ste. 400, Dallas, TX 75247) and Jack E. Randorff (Randorff and Assoc., Inc., Ransom Canyon, TX 79366)

Noise intrusions from the Miller Outdoor Theater in Houston, Texas caused recurring complaints from the surrounding upscale community. Measurements confirmed that events at this performing arts venue regularly produced levels of 75 dBA and above at residents homes approximately 1,300 ft from the proscenium stage. The project goals were to provide the theater with a high-quality sound system, capable of producing nominal levels of 95 dBA at the audio mixing console and 85 dBA at the farthest lawn seats (340 ft from the stage) while not exceeding levels of around 55 dBA in the nearby neighborhood. These goals were met through loudspeaker selection (including some customization) along with careful loudspeaker zoning, elevation, and aiming. Two different computer-based programs were used for acoustical modeling—one to design and predict the sound systems coverage throughout the seating area and another to predict noise propagation from the theater into the community. Both programs were used to produce maps indicating sound levels as color contours which simplified presentation of these data to the client.

2:55

3pAA6. The acoustics of the outdoor Shakespeare theater—Rutland, UK. Peter A. Mapp (Peter Mapp Assoc., Colchester C03 4JZ, UK)

The outdoor Shakespeare Amphitheatre at Tolethorpe (Rutland, UK) is an unusual form of theater in that the actors perform in the open air while the audience is enclosed in a lightweight Teflon fabric auditorium. This unique format produces some unusual acoustic properties and problems. The early reflection patterns from and around the stage are quite different from a normal enclosed proscenium theater. The Teflon walls and roof, while effectively acoustically transparent at low and lower midfrequencies, become highly reflective at high frequencies. This not only impacts the transmission of sound through the structure but also speech transmission and intelligibility within the auditorium. The hyperbolic curved-roof surfaces and flat parallel walls were found to cause undesirable sound focusing and reflections detrimental to speech intelligibility. The acoustics of the space was investigated by means of both directional and binaural time-domain spectrometry techniques as well as more traditional acoustic measures. The paper presents the results of these measurements and outlines the remedial treatments adopted.

WEDNESDAY AFTERNOON, 6 DECEMBER 2000 PACIFIC SALONS A–F, 1:55 TO 3:00 P.M.

Session 3pID

Interdisciplinary: Hot Topics in Acoustics

Ralph T. Muehleisen, Chair
Civil, Environmental and Architectural Engineering, University of Colorado, Campus Box 428, Boulder, Colorado 80309-0428

Chair's Introduction—1:55

Invited Papers

2:00

3pID1. Hot topics in acoustical oceanography. James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Acoustics continues to play an ever-increasing role as a means of exploring the ocean. In every area of oceanography, acoustic techniques are being used to make measurements that would have been considered science fiction only a few scant years ago. In this lecture, we will show some of the recent, sometimes startling advances that have been made in oceanography using acoustics. We will also endeavor to put the acoustics measurements in the context of other types of measurements that are made. Starting with physical oceanography, we will look at the advances that have been made, going from the largest (gyre) scales to the smallest (microscopic turbulence) oceanographic scales. We will then turn to marine geology, and examine how various acoustic techniques both map the structure of the marine sediments and also help us understand the formation of sedimentary structures. Finally, we will examine marine biology, where acoustic techniques are helping us quantify the abundance of marine life and, at a more basic level, understand marine ecosystems.

2:20

3pID2. Recent developments in biomedical ultrasound. Lawrence A. Crum (Appl. Phys. Lab., 1013 NE 40th St., Univ. of Washington, Seattle, WA 98105)

Biomedical ultrasound has seen remarkable advances in recent years. By utilizing the properties of nonlinear acoustics, diagnostic ultrasound has shown increased applicability to a wide number of clinical conditions and pathologies. Techniques such as harmonic imaging and the use of ultrasound contrast agents (stabilized microbubbles) have enabled such long-sought goals as noninvasive determination of myocardial perfusion to be clearly within our grasp. Advancements in semiconductor miniaturization have led to the construction of ultrasonic scanners that are now hand-held, and together with telemedicine techniques, it is now reasonable to expect that diagnostic ultrasound will soon be the doctor's stethoscope. An even more promising future is seen for therapeutic ultrasound. Although the mechanism is not yet clearly understood, ultrasound can transiently permeabilize cell membranes, thus permitting the delivery of therapy to specific sites within the body; indeed, together with drug-carrying ultrasound contrast agents, ''site-specific drug delivery'' is now in clinical trials. Finally, the application of High Intensity Focused Ultrasound can induce coagulative necrosis at well-controlled sites within tissue. When imaging and therapy are combined, ''image-guided, transcutaneous, bloodless surgery'' devices are now under development. With acoustics, ''Star Trek medicine'' is just around the corner.

2:40

3pID3. Recent studies of the lip reed: Low-tech experiments and stroboscopic observations. R. Dean Ayers (Dept. of Phys. and Astron., California State Univ., Long Beach, 1250 Bellflower Blvd., Long Beach, CA 90840, rdayers@csulb.edu)

There are currently three basic models for the brass player's vibrating lips: (1) Helmholtz's outward-striking or swinging door model, (2) a sliding door model, with motion that is transverse to the air stream, and (3) a hybrid, two-dimensional model that combines those two motions. Time-domain computer simulations of these models have been carried out by S. Adachi and M. Sato [J. Acoust. Soc. Am. **97**, 3850–3861 (1995); **99**, 1200–1209 (1996)]. The experience of brass players agrees qualitatively with the behavior of the hybrid model. Feedback is optimized near each peak of the input impedance curve, and from there it is easy to drop the playing frequency but not to push it upward. (Buzzing on an isolated mouthpiece yields the same behavior, even though there is no regime of oscillation.) A graph of playing frequency versus an operationally defined ''lip frequency'' is obtained by using a side hole in the backbore. Agreement with results from the hybrid model is good for the first few modes but not the higher ones. Stroboscopic images of Rayleigh waves on the upper lip suggest that a more advanced model should have at least two masses performing orbital motion. Demonstrations will be included. [Work supported in part by the Scholarly and Creative Activities Committee at CSULB.]

WEDNESDAY AFTERNOON, 6 DECEMBER 2000 YAWL ROOM, 1:00 TO 2:35 P.M.

Session 3pMU

Musical Acoustics: Asian Musical Instruments and Traditions II

Thomas D. Rossing, Chair
Physics Department, Northern Illinois University, DeKalb, Illinois 60115

Chair's Introduction—1:00

Invited Papers

1:05

3pMU1. Acoustics of Asian drums. Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115)

Drums have played an important role in nearly all musical cultures, including those of Asia. The tabla and mrdanga of India are specially tuned to have harmonic partials so that they convey a strong sense of pitch. The very important drums in Balinese and Javanese gamelans, on the other hand, do not have a strong pitch. The large o-daiko of Japan have modes of vibration that are somewhat similar to those in the bass drum. Smaller drums, such as the tsuzumi, have spectra that vary considerably with playing style.

1:35

3pMU2. Tuned sets of bells from ancient China. Robert W. Bagley (Dept. of Art & Archaeology, Princeton Univ., 105 McCormick Hall, Princeton, NJ 08544-1018, rwbagley@princeton.edu)

Tuned sets of bronze bells appeared very early in China, sometime around 1200 B.C., and over the next millennium underwent an extraordinary development. The most spectacular set known so far consists of 65 bells found in 1978 in the tomb of Marquis Yi of Zeng (deceased 433 B.C.), ruler of a small and previously unknown state in central China. The set has a range of five octaves, within which is a three-octave chromatic stretch; the largest bell weighs 200 kg. Despite much study of this and other sets, it remains unclear

how tuned sets were designed and cast; it has not been possible to deduce a scaling rule from measurements of bells. Invoking both archaeological evidence and musical considerations, a hypothesis will be presented to suggest how tuned sets might have been invented and elaborated.

2:05

3pMU3. Sounds of ancient Korean bells. Yang-Hann Kim (Ctr. for Noise and Vib. Control (NOVIC), Dept. of Mech. Eng., KAIST, 373-1 Science Town, Taejon-shi, Korea)

This paper reviews the sounds of bells cast during the Shilla Dynasty. One unique characteristic of Korean bells is that they create beats. The beats of various ancient bells are summarized and analyzed with regard to mass, and other geometrical dimensions. Other interesting and unique features of the old Korean bells are that they have sound pipes on top and hollows below the bells. The types of pipes and hollows are also studied and reported. In fact, the pipe is a high-pass filter so that it dissipates high-frequency sound as fast as it can, and the hollow is a resonator that can somehow sustain a beating sound. Bells cast during the Shilla and Korea Dynasty and their sound characteristics are investigated and summarized.

WEDNESDAY AFTERNOON, 6 DECEMBER 2000 CALIFORNIA SALON 3, 2:00 TO 3:05 P.M.

Session 3pNS

Noise and Archives and History: Acoustical Society's Role in Noise Control

Tor S. D. Kihlman, Chair
Applied Acoustics, Chalmers University of Technology, Goteborg S-412 96, Sweden

Chair's Introduction—2:00

Invited Paper

2:05

3pNSa1. The Acoustical Society of America's historical contributions to noise and its control. Leo L. Beranek (975 Memorial Dr., Ste. 804, Cambridge, MA 02138-5755, beranekleo@mediaone.net) and William W. Lang (29 Hornbeck Ridge, Poughkeepsie, NY 12603-4205)

The Society's contributions to noise and its control, a subfield of the broad subject of acoustics, have been significant and constant over the years since the first noise paper was presented at the Society's first meeting in 1929. The major contributions are considered in an historical context—during the period before World War II when the Society first produced national standards for instrumentation used for noise measurement, during the War and the Postwar Era when there was considerable government support for airborne acoustics research, during the 1950s when *Noise Control* magazine was published, followed by the 1960s and the later years of the 20th Century. Of the ASA members expressing interest in the 13 different subfields of acoustics, the number concerned with noise is near the top of the list. But ASA is no longer alone. Nearly a dozen other professional organizations in the U.S.A. have groups that are involved with noise and its control. Over the past three-quarters of a century, ASA has played a leadership role in developing the subfield of acoustics focused on noise control, and in producing a broad range of American National Standards on noise and vibration. In the future, ASA must take concrete, positive steps to retain its position of leadership.

A NOTE ABOUT THE ASA HISTORY LECTURE SERIES

In 1997, the ASA Committee on Archives and History conceived a plan for a series of invited lectures on each of the technical areas of the Society which would memorialize the significant achievements and milestones of each of its twelve technical committees and one interdisciplinary technical group during the first three quarters of the Society's first century.

With the cooperation of the technical committees, distinguished individuals are selected to review the history of their particular technical speciality and present a lecture which shows how that activity has developed and has contributed to the Society at large and to the broad field of acoustics as well. At the meetings in Columbus, Ohio and Atlanta, Georgia, the first four History Lectures were presented: Gabriel Weinreich on Musical Acoustics, Robert Beyer and David Blackstock on Physical Acoustics, Stanley Ehrlich on Engineering Acoustics and Ewart Wetherill on Architectural Acoustics. At each subsequent meeting two additional lectures will be scheduled including those in Noise and Speech Communication at this meeting in Newport Beach.

The invited lecturers have been asked to prepare a written manuscript of their lectures which will be published in a commemorative book for the 75th Anniversary of the Society to be celebrated in 2004. The Archives and History Committee and the individual technical committees/group welcome comments and suggestions on both the History Lecture Series and on the proposed ASA Diamond Anniversary Book. Volunteers to assist the committees would be most welcome too. Contact Henry Bass, Chair, Committee on Archives and History, pabass@sunset.backbone.olemiss.edu

Session 3pPA

Physical Acoustics: Thermoacoustics II

Larry A. Wilen, Chair
Department of Physics and Astronomy, Ohio University, Clippinger Laboratories, Athens, Ohio 45701-0882

Chair's Introduction—1:40

Contributed Papers

1:45

3pPA1. Time-averaged pressure drop produced by an abrupt change in a resonator's cross section. A. Doller, Anthony A. Atchley, and Roger Waxler (Grad. Prog. in Acoust., The Penn State Univ., State College, PA 16804, aatchley@psu.edu)

The behavior of steady, incompressible flows through constrictions, expansions, bends, and branches in pipes is very well documented and commonly used to design fluid handling systems over a wide range of Reynolds numbers, surface roughness, etc. However, the state of knowledge of high-amplitude acoustic flows in similar structures is much less mature, as demonstrated by recent developments in thermoacoustic devices [S. Backhaus and G. W. Swift, J. Acoust. Soc. Am. **107**, 3148–3166 (2000)]. The research presented here, motivated by this lack of understanding, focuses on the nature of predominately reactive, high-amplitude acoustic fields in the vicinity of changes in resonator cross section. Oberst-style resonators [H. Oberst, Akust. Z. **5**, 27–38 (1940)], consisting of two sections of straight brass pipe of different diameters and joined through either step-like or conical couplers, were constructed. The resonators are driven by a graphite piston at one end, and rigidly terminated at the other. The acoustic pressure is measured at four locations along the resonator. Of particular interest is the time-averaged pressure drop across the coupler section. Measurements of this pressure drop are reported as functions of drive amplitude for the different coupling configurations. [Work supported by the Office of Naval Research.]

2:00

3pPA2. Practical considerations for the design and construction of an acoustic dynamometer. Robert W. M. Smith, John F. Heake, and Steven L. Garrett (Penn State Appl. Res. Lab, P.O. Box 30, State College, PA 16804)

An instrumented, variable acoustic load of the type described by Fusco [Fusco *et al.*, J. Acoust. Soc. Am. **91**(4), 2229–2235 (1992)] and Gardner [D. L. Gardner and G. W. Swift, Cryogenics **37**(2), 117–121 (1997)] has been implemented for the testing of high power (greater than 1.5 kW) electrodynamic loudspeakers for thermoacoustic refrigerator applications. The load serves to facilitate performance measurements for these loudspeakers under various operating conditions, a role analogous to a dynamometer in conventional rotating machinery. Attached to a primary acoustic resonator, the acoustic dynamometer consists of a 2-in angle globe valve and other flow constrictions connected to an acoustic compliance. The design of these flow restrictions employed the acoustic ''minor loss'' model developed by Swift *et al.* [Swift *et al.*, J. Acoust. Soc. Am. **105**, 711–723 (1999)]. Measurements of the dynamometer performance agreed with the calculated acoustic impedance to within 16%. At powers exceeding 1 kW, agreement between the acoustic load dissipation and other measurements of loudspeaker power agree to within 3%. To illustrate the utility of an acoustic dynamometer, a brief summary of measurements on two moving-magnet loudspeakers will be presented. [Work supported by ONR.]

2:15

3pPA3. Wet-walled thermoacoustics. William V. Slaton and Richard Raspet (Dept. of Phys. and Astron. and The Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677, wmslaton@meta3.net)

An analytic solution of sound propagation in wet-walled tubes with a temperature gradient will be presented. The tube contains an inert gas-vapor mixture with a thin layer of condensed vapor coating the tube wall. The vapor phase condenses and evaporates from this layer during an acoustic cycle. This phased evaporation and condensation modifies traditional energy density and wave number equations. It is found that mass and heat transport act in parallel thereby increasing the acoustic energy density for proper choice of gas mixtures. This increase in energy density due to inert gas-vapor working fluids can be quite large and even dominate the traditional thermal transport contribution. The use of phased evaporation-condensation in gas mixtures provides a path to higher energy density in thermoacoustic systems. The effects on sound propagation, energy density, and efficiency due to the presence of an inert gas-vapor working fluid in tubes with wet walls will be discussed. [Work supported by the Office of Naval Research.]

2:30

3pPA4. Torsionally resonant toroidal standing-wave resonator. Scott R. Diehl, Steven L. Garrett (Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804), Milo Friesen, Roy L. Kessinger, Jr., Paul Receveur, and Leonard Schrank (Lynx Motion Technol. Corp., New Albany, IN 47151-0408)

A toroidal resonator with a rigid barrier can be used to generate high-amplitude standing waves if the resonator is oscillated in a torsional mode about its central axis [S. Garrett, US Pat. No. 5,953,921 (Sept. 21, 1999)]. In the following, a 2-1/2-in diameter schedule-40 steel pipe is bent into a 13-in radius torus to form the resonator. The torus is then deformed along its cross section to detune harmonic overtones and suppress shock wave formation. The torsional drive force is provided by rigidly attaching the resonator to the housing of a modified Lynx Model T468 Direct-Drive Servo Motor [Kessinger *et al.*, US Pat. No. 5,744,896 (April 28, 1998)]. The resonator/motor housing and shaft/rotor assemblies are made into a mechanical 2-DOF torsional oscillator by coupling their moments-of-inertia through an elastic restoring force. The torsional restoring force is provided by a quartet of semi-cylindrical stainless steel (17-7 PH) springs mounted between the resonator/housing and shaft/rotor assemblies. Measurements of modal anharmonicity and peak acoustic pressures at resonance will be reported and compared to DeltaE model results. [Work supported by the Office of Naval Research.]

2:45

3pPA5. Nostack thermoacoustics. Ray Scott Wakeland and Robert M. Keolian (Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804, wakeland@psu.edu)

The ''stack'' used in conventional standing-wave thermoacoustics depends on intrinsic thermodynamic irreversibilities for its operation. One approach to increasing efficiency is to eliminate the stack, retaining a small gap between heat exchangers in a standing wave. While a ''nos-

tack'' device has a temperature span that is limited by the achievable pressure amplitude, it is expected to have greatly increased thermodynamic efficiency compared to a stack-based device with the same power and operating temperatures. Losses associated with flow through the heat exchangers, however, are expected to be higher in a nostack device. The standard theory of low-amplitude thermoacoustics does not apply to nostack devices, where the displacement amplitude is larger than the heat exchangers and the gap. As an alternative, an idealized thermodynamic model of nostack is combined with separate calculations of conduction loss and flow losses, including so-called ''minor losses'' arising from sudden changes in cross-sectional flow area. The results suggest that by positioning the heat exchangers very close to a pressure antinode to reduce flow velocity, efficiencies similar to Stirling-like thermoacoustic devices might be achieved with the nostack idea. [Work supported by the Office of Naval Research and the Pennsylvania Space Grant Consortium.]

3:00

3pPA6. Alternative-geometry traveling-wave thermoacoustic devices. Larry Wilen and Gabriela Petculescu (Dept. of Phys. and Astron., Clippinger Lab, Ohio Univ., Athens, OH 45701, wilen@helios.phy.ohiou.edu)

Ceperley has discussed the idea that an acoustic traveling wave passing through a regenerator executes a Stirling cycle [P. H. Ceperley, J. Acoust. Soc. Am. **77**, 1239–1244 (1985)]. He also emphasized that for an acoustic Stirling device to be useful, the impedance must be high or else viscous losses will dominate. Backhaus and Swift recently found a way to produce a useful device with high impedance and traveling wave phasing using a toroidal geometry [S. Backhaus and G. W. Swift, J. Acoust. Soc. Am. **107**, 3148–3166 (2000)]. We will discuss possible ways to realize these conditions with alternative geometries, and we will present some preliminary results for a refrigerator employing one such configuration. [Work funded by the Office of Naval Research.]

WEDNESDAY AFTERNOON, 6 DECEMBER 2000

CATAMARAN/TRIMARAN ROOMS, 1:00 TO 3:00 P.M.

Session 3pPP

Psychological and Physiological Acoustics: Potpourri (Poster Session)

Thomas Z. Strybel, Chair

Psychology Department, California State University at Long Beach, 1250 Bellflower Boulevard, Long Beach, California 90840

Contributed Papers

All posters will be on display from 1:00 p.m. to 3:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:00 p.m. to 2:00 p.m. and contributors of even-numbered papers will be at their posters from 2:00 p.m. to 3:00 p.m.

3pPP1. Effects of basilar membrane roughness and nonlinearity on stimulus frequency otoacoustic emissions. Carrick L. Talmadge (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677) and Arnold Tubis (Purdue Univ., West Lafayette, IN 47907)

Stimulus frequency otoacoustic emissions (SFOAEs) may be evoked via basilar membrane (BM) nonlinearity, BM roughness (random inhomogeneities), or a combination of these properties. Emissions generated by BM nonlinearity alone have phase gradients that are too small to account for the observed emission fine structure [e.g., Shera and Guinan (1999)]. Linear reflection via BM roughness can account for the main phase characteristics of the emissions [e.g., Shera and Zweig (1993); Talmadge *et al.* (1995)]. It is shown theoretically that, contrary to what is expected from naive considerations concerning phase derivatives and group delays, the physical time delays associated with both emission mechanisms are approximately equal to the round-trip cochlear wave travel time from the cochlear base to the tonotopic place for the stimulus frequency. A theoretical framework is presented for describing the combined effects of roughness and nonlinearity on the emission fine structure. For the case of cochlear roughness and a weak nonlinearity, it is shown that in some instances, the nonlinearity can significantly modify the level and phase behaviors of the fine structure and actually enhance the fine structure patterns. The implications of these results for the stimulus level dependence of SFOAE fine structure are also explored.

3pPP2. Hydrodynamics of otoacoustic emissions. Taha Jaffer and Hans Kunov (Inst. for Biomaterials and Biomed. Eng., Univ. of Toronto, 4 Taddlecreek Rd., Toronto, ON L4J 2X7, Canada, taha@ibme.utoronto.ca)

In this paper, a novel boundary condition for the classical cochlear model will be outlined that allows the realistic simulation of tone-burst transiently evoked otoacoustic emissions (TEOAEs). OAEs are sounds that are emitted from the inner ear and have been related to hearing function. Classical cochlear models suffer from an inability to realistically simulate OAEs because of the simplistic treatment of the hydrodynamics at the interface between the oval window, stapes, and cochlear fluid. In the classical cochlear model, there are two boundary conditions; the first describes the helicotrema, and the second details how fluid is displaced by the oval window in response to sound entering the ear. In this model, OAE production cannot be simulated because once the stimulating sound is stopped, the oval window is immobile and incapable of transmitting acoustical energy back out of the inner ear. In this paper, this boundary condition will be reconsidered in light of the additional fluid displacement caused by the movement of the cochlear partition. The addition of this boundary condition yields realistic simulations of tone-burst OAEs, both in time evolution and frequency content.

3pPP3. Modulated otoacoustic emissions reveal features of the human cochlear impedance at low frequencies. Torsten Marquardt (Dept. of Physiol., Univ. College London, Gower St., London WC1E6BT, UK, torsten.marquardt@ucl.ac.uk), Johannes Hensel, Guenther Scholz, and Dieter Mrowinski (Virchow Hospital, Augustenburger Platz 1, 13353 Berlin, Germany)

Basilar membrane (BM) displacement by loud low-frequency (LF) tones periodically alters the level of distortion-product otoacoustic emissions (DPOAEs) as a function of the phase of the modulating LF tone. For three human subjects the level of a modulating LF tone (15–320 Hz) was adjusted to maintain a constant DPOAE-modulation depth, indicating a constant periodic LF BM displacement. The resulting modulation LF-tone levels match the iso-loudness contours (ISO/R226) except for an antiresonance centered at about 55 Hz. Here, an increase in level of approximately 5 dB above the iso-loudness contour is required to maintain constant periodic LF BM displacement. The antiresonance separates two distinct regions of the cochlear impedance: a slope of 12 dB/octave at frequencies below 55 Hz suggests a mass-controlled impedance resulting from perilymph flow through the helicotrema; a slope of 6 dB/octave above 55 Hz suggests the existence of a traveling wave with resistive impedance. Psychophysical threshold experiments did not show the antiresonance. Antiresonances have been observed previously in animal experiments using cochlear microphonics or perilymph pressure measurements to estimate BM displacement. Modulated DPOAEs are a noninvasive method to investigate the BM movement at low frequencies.

3pPP4. Effect of tonal duration on auditory evoked potentials as a function of age. Jodi Ostroff, Kelly McDonald, Claude Alain (Rotman Res. Inst., Baycrest Ctr. for Geriatric Care, Univ. of Toronto, 3560 Bathurst St., Toronto, ON M6A 2E1, Canada), and Bruce Schneider (Univ. of Toronto, Mississauga, ON L5L 1C6, Canada)

Behavioral research has shown that temporal resolution decreases with age as evidenced by higher-gap-detection thresholds in older versus young adults with normal hearing. In the present study, we examined the neural correlates of the temporal resolution function in young, middle aged, and older subjects, using a brief Gaussian-shaped 2-kHz tone that increased by 2 ms in separate trial blocks. All stimuli elicited a negative potential that peaked at about 100-ms poststimulus (N1). The N1 amplitude increased linearly with increases in tonal duration. The growth function of the N1 amplitude was similar across the three age groups, although the slope of the function in older adults was shallower. Results suggest that the N1 response can resolve temporal information on the order of 2 ms. This electrophysiologic threshold parallels behavioral findings. The slower growth function of the N1 response in older adults may reflect a deficit in temporal processing. [Work supported by the Medical Research Council of Canada.]

3pPP5. The form of the expansive nonlinearity in cochlear implant stimulation. Monita Chatterjee (Dept. of Auditory Implants and Percept., House Ear Inst., 2100 W. Third St., Los Angeles, CA 90057)

In electrical stimulation with cochlear implants, loudness grows as an expansive function of stimulus amplitude. The mathematical form of the nonlinearity is difficult to discern, primarily because of the limited current range over which measurements can be made without reaching uncomfortably loud sensations. Both power and exponential functions appear to fit measured loudness estimates reasonably well over the dynamic range. For a given pulse train, increasing the pulse phase duration (D) and/or the separation between active and return electrodes in the cochlear (M) results in increased loudness. Power function fits to the loudness functions show that while the exponent of the power function is insensitive to these stimulus parameters, the multiplicative constant is not. In contrast, the best fitting exponential functions show systematic changes in the exponent with changing M and D. The two kinds of functions also yield different predictions for the level-dependence of modulation sensitivity. When the nonlinearity is a power function (as in normal hearing), modulation sensitivity is predicted to show negligible dependence on level; when the nonlinearity is exponential, a strong level-dependence is predicted. In cochlear implant listeners, modulation sensitivity is strongly level-dependent, tipping the balance in favor of an exponential nonlinearity. [Work supported by NIDCD.]

3pPP6. Speech recognition under conditions of frequency-place expansion and compression. Deniz Baskent (Dept. of Biomed. Eng., Univ. of Southern California, Los Angeles, CA 90089) and Robert Shannon (House Ear Inst., Los Angeles, CA 90057)

The mapping of frequency information onto the correct cochlear place is critical for speech recognition. In cochlear implants, the cochlear tonotopic range represented by the electrode array is smaller than the acoustic frequency range used in the speech processor. While this condition utilizes a wide range of speech frequency information, it results in a compression of the tonotopic pattern of speech information delivered to the brain. An alternative approach is to take the core frequency range necessary for speech and expand its representation in the cochlea, like an ''acoustic fovea.'' This study examined the effect of linear frequency-place compression and expansion on speech intelligibility for various cochlear locations and number of spectral channels. These conditions were presented to normal-hearing listeners using a noise band vocoder. The cochlear tonotopic range was held constant by employing the same carrier bands for each condition, while the frequency range of the analysis bands was changed. For each condition, the result was compared to that of the perfect tonotopic match, where the carrier and the analysis bands were perfectly matched. Speech recognition in frequency-place expansion and compression was always equal to or poorer than the matched condition. [Work supported by NIDCD.]

3pPP7. Horizontal and vertical minimum audible angle: Effects of stimulus onset asynchrony and azimuth. Thomas Z. Strybel, Eishi Adachi, and Simon Jeon (Dept. of Psych., California State Univ., Long Beach, Long Beach, CA 90840)

The effects of stimulus onset asynchrony (SOA, the onset-onset time difference) on the minimum audible angle (MAA) were measured for horizontal and vertical orientations and two azimuth locations. Each participant was tested in 16 conditions created by all combinations of four SOAs (25, 50, 100 and 400 ms), two orientations (horizontal and vertical), and two azimuth locations (0 and 90 deg). MAAs were measured with an adaptive procedure. The stimulus was high pass noise with a lower frequency cutoff of 4 kHz and amplitude of 55 dB A-weighted. The burst duration was 50 ms with a rise/decay time of less than 1 ms. As expected, horizontal MAAs were lowest at 0 deg and increased substantially at 90 deg; vertical MAAs were only minimally affected by azimuth. The effect of SOA was similar for both orientations and azimuths. In all conditions, the MAA decreased exponentially with SOA. The slope of the horizontal function at 0 deg was slightly steeper than the slopes of the remaining conditions.

3pPP8. Objective discrimination of extensity: Concurrent versus successive sound sources. Prisilia Tirtabudi and David Perrott (Dept. of Psych., California State Univ., Los Angeles, 5151 State University Dr., Los Angeles, CA 90032, tirtabudi@hotmail.com)

This study compared human ability in processing and detecting concurrent versus successive sound stimuli and the degrees of angular separation between sources. Bursts of broadband white noise at 85 dB (SPL) were presented with duration of either 1000, 300, 100, 30, or 10 ms. Angles between speakers for each duration time were at 4.5°, 9.0°, 18°, 27°, and 36°. Two pairs of bursts were presented at a time, either simultaneously or sequentially. Participants were to discriminate whether the first pair of bursts had a larger angular separation between sources than the second pair. Listeners' performance were almost similar between concur-

rent versus successive stimuli when the durations were 1000, 300, and 100 ms. However, as the duration of the stimuli dropped to 30 and 10 ms, so did the performance in the concurrent condition. In addition, as the angles between the two speakers grew farther apart in both conditions, the degrees of jnd (MAA) were also increasing. Results will be discussed in terms of the auditory system's ability to resolve spatially distributed sources that are set off either concurrently or sequentially.

3pPP9. Psychophysical tuning curves and filter shapes sharpen over time. Jennifer L. Repovsch and Sid P. Bacon (Psychoacoustics Lab., Dept. of Speech and Hearing Sci., Arizona State Univ., P.O. Box 87-1908, Tempe, AZ 85287-1908, spb@asu.edu)

Psychophysical tuning curves (PTCs) measured in simultaneous masking usually sharpen as a short signal is moved from the onset to the temporal center of a longer duration masker. Filter shapes derived from notched-noise maskers have not consistently shown this effect. One possible explanation for this difference is that the PTC paradigm uses a fixed signal level, whereas filter shapes typically have been derived from results where the noise masker was fixed in level. In the present study, PTCs and filter shapes were measured with a signal fixed in level at 10 dB SL. The masker was either a pure tone (PTCs) or a noise with a spectral notch placed symmetrically or asymmetricaly about the 2-kHz signal frequency (filter shapes). The 20-ms signal was presented at the onset or temporal center of the 400-ms masker. For a given signal delay, the PTCs were substantially sharper than the filter shapes. Importantly, however, for all four subjects, the PTCs and the filter shapes sharpened on both the low- and high-frequency sides as the signal moved from the beginning to the temporal center of the masker. This sharpening was considerably greater for the filter shapes. [Work supported by NIDCD.]

3pPP10. Up/down asymmetry in FM identification. Michael Gordon and David Poeppel (Dept. of Biol., CEBH, Dept. of Linguist., Cognit. Neurosci. of Lang. Lab., Univ. of Maryland, College Park, MD 20742)

Temporal limits of the human auditory system with regard to identifying the direction of frequency modulation (FM) were investigated: How brief can frequency change be without compromising FM direction identification? Such abilities may relate to processing formant transitions. In a 2AFC paradigm, subjects binaurally heard linear FM sweeps rising or falling in frequency between 1000 and 1500 Hz. Ten FM rates were used for both FM directions. The FM rate was modulated by varying stimulus duration from 5 to 640 ms. Each FM stimulus was presented 20 times, in pseudorandom order, for a total of 400 trials. Subjects ($n=24$) identified the direction of frequency change. There was a marked difference between up and downward FMs. For upward FMs, 90% accuracy in direction identification was achieved at a rate as fast as 17 Hz/ms. A considerably slower rate, 3 Hz/ms, was needed for identifying downward FMs with an accuracy of 90%. Humans are better at identifying rapid upward FMs than downward ones. This result is congruous with studies revealing lower thresholds for detecting rapid upward FMs (Collins and Cullen, 1978) and higher magnitude evoked potentials elicited by upward FMs (Maiste and Picton, 1989). [Work supported by Comparative and Evolutionary Biology of Hearing Training Program.]

3pPP11. Greek bilingual listeners perceive the tritone stimuli differently from speakers of English. Magdalene H. Chalikia, Angeline M. Norberg (Psych. Dept., Minnesota State Univ., Moorhead, MN 56563), and Lena Paterakis (The American College of Greece, Aghia Paraskevi, Greece)

When listeners are presented with pairs of octave-complex tones related by a tritone interval (a half-octave), they hear the pattern as ascending or descending, according to an individual pitch class template. Deutsch (1991) has claimed that this template may be influenced by language. In order to test this hypothesis, data from Greek bilingual listeners were collected and compared with data from Texas, California, and the south of England. The results show significant differences in how Greek listeners hear the tritone stimuli, as compared to listeners in the other groups. There is also evidence that the Greek listeners may have developed two different pitch class templates, possibly representing the influence of English and the influence of Greek.

3pPP12. Listeners in Sweden perceive tritone stimuli in a manner different from that of Americans and similar to that of British listeners. Magdalene H. Chalikia and Fredrik Leinfelt (Dept. of Psych., Minnesota State Univ., Moorhead, MN 56563)

The tritone stimuli consist of two tones, each of which has octave-related harmonics. The two tones are separated from each other by a half-octave (a tritone interval). The tones are typically presented successively, and listeners decide whether or not the second tone in the pair is perceived as ascending or descending, relative to the first one. Responses tend to vary across listeners for the same stimuli, and for different stimuli. In this study, participants from Stockholm, Sweden, were presented with such tritone pairs and were asked to determine if they heard each pair as ascending or descending. Deutsch's (1991) hypothesis, that language background may influence the perception of the tritone stimuli, was tested. All participants were speakers of both English and Swedish. Their data were compared to data of English monolinguals tested in Texas, and California, as well as listeners from the south of England. Preliminary results suggest that the overall perception pattern in Sweden differs from that found in Texas and California, but tends to be similar to the perception pattern found with English listeners.

3pPP13. The perception of tritone stimuli by listeners in the Midwest. Meredith Haugen and Magdalene H. Chalikia (Psych. Dept., Minnesota State Univ., Moorhead, Moorhead, MN 56563)

The tritone stimuli [Deutsch (1986)], consist of two tones, each with octave-related harmonics. The tones are separated from each other by a tritone (half-octave) interval. The tones are presented successively, and listeners decide whether or not the second tone in the pair is perceived as ascending or descending, relative to the first one. Listeners tend to hear the tritone pattern as ascending or descending, according to an individual pitch class template, which Deutsch (1991) claims may be influenced by language or English dialect. Some similarities have been found among people from different areas of the U.S., suggesting a canonical manner of perceiving the tritone stimuli, probably propagated by the media [Ragozzine and Deutsch (1994)]. The present study tested the media exposure hypothesis, by presenting tritone stimuli to listeners who grew up in the Midwest. Results show that Midwesterners show a different perceptual pattern relative to that of listeners from Texas and the south of England, but the same pattern with that heard by listeners from California. Spectral envelope effects were also examined, and suggest that the perceptual patterns found under high and low envelopes are similar to those found for Californians, but different from those found with listeners from Texas.

3pPP14. Investigation of the influence of visual stimuli on the breakdown of the echo threshold. Thomas Djelani (Institut fuer Kommunikationsakustik, Ruhr-Universitaet Bochum, 44780 Bochum, Germany)

Audiovisual interaction is already known from the ventriloquism effect [W. R. Thurlow and Ch. E. Jack, Percept. Mot. Skills **36**, 1171–1174 (1973)]. Synchronously presented auditory and visual stimuli are perceptually grouped and localized as one object. The performed experiments aimed to explore a possible influence of visual stimuli on echo suppression. They focused on investigating whether visual stimuli can also cause a breakdown of the echo threshold. One hypothesis is that the breakdown of echo suppression is caused by violations of the listeners' expectations

concerning the presented scenario. The question is whether this implausibility can be attributed to acoustical reasons. Two different kinds of visual implausibilities were tested by measuring echo threshold with an adaptive paradigm. In the build-up phase of both experiments, visual and auditory parts of the stimuli have been presented from the same direction. In the test phase, the visual stimulus' direction was altered. The main result of the investigations was that both experiments did not provide any evidence of a visually induced breakdown of the echo threshold. This holds true, even if the position of the auditory event corresponding to the direct sound was changed due to the ventriloquism effect by changing the corresponding visual stimulus' direction of incidence.

3pPP15. Monitoring and localizing simultaneous real-world sounds: Implications for the design of spatial audio displays. Diane L. Guettler (Dept. of Psych., California State Univ., Long Beach, 1250 Bellflower Blvd., Long Beach, CA 90840), Robert S. Bolia (Air Force Res. Lab., Wright–Patterson Air Force Base), and W. Todd Nelson (marchFIRST)

Numerous researchers have investigated the extent to which human listeners are able to localize free-field and virtual acoustic stimuli. Despite the abundance of this research, few studies have focused on the number of sounds a listener is able to localize simultaneously, and even fewer have employed naturally occurring or real-world sounds. Results from these types of experiments may have important implications at the basic research level within the fields of perception and attention, and may also prove useful for guiding the design of spatial audio displays in applied disciplines such as engineering and human factors. In the present study, listeners localized between one and three virtual acoustic stimuli, the positions of which were restricted to the horizontal plane under factorial combinations of seven spatial locations, three set sizes (one, two, or three sources), and three distinct acoustical sounds of a real-world nature (e.g., a crying baby). The paradigm employed required participants to monitor a set of simultaneously presented sounds and localize one specific sound when prompted. Performance was measured in terms of average localization error and percentages of front–back confusions. These results will be discussed in terms of their implication for the design of spatial audio interfaces for complex, multitask environments.

WEDNESDAY AFTERNOON, 6 DECEMBER 2000 CALIFORNIA SALON 5, 1:30 TO 2:45 P.M.

Session 3pSA

Structural Acoustics and Vibration: Acoustic Scattering from Elastic Structures

J. Gregory McDaniel, Chair

Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02537-0339

Contributed Papers

1:30

3pSA1. Scattering from a partially filled elastic-shelled sphere. John A. Fawcett (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada)

Scattering from spherical scatterers can be analytically formulated in terms of a harmonic series solution. The case of a shelled sphere with an evacuated or filled interior has been particularly well studied. In this paper, the case of a partially filled sphere is considered. A solution for the scattered field is sought in terms of a harmonic Legendre function series. A standard series expansion is used for the shell and the exterior of the sphere. A modified expansion, which satisfies the interface condition within the interior, is used for the inner region. Unlike the standard spherical scattering problem, the system of equations for the various orders (for a fixed azimuthal order) of the Legendre functions are now coupled. Numerical computations are presented for a steel-shelled sphere with various levels of water-fill.

1:45

3pSA2. Causal and unique reconstruction of nonminimum reflection phase for a class of submerged structures. Cory L. Clarke and J. Gregory McDaniel (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215)

An approach is presented for uniquely determining the phase of an acoustic reflection from an object whose impedance magnitude approximates the acoustic impedance of the surrounding medium. This approximation defines a semiannular region in the complex impedance plane that lies between the rigid and pressure-release limits. When this region is mapped onto the complex reflection coefficient plane, the approximation allows for a wide range of reflection coefficients. A Taylor series expansion for the square of the reflection coefficient magnitude is used to estimate the phase of the reflecting impedance. The causality condition is invoked to recover the magnitude of the impedance from its phase. Once the magnitude and phase of the impedance is known, the complex-valued reflection coefficient algebraically follows regardless of whether or not the reflection coefficient is minimum phase. The approach is useful in experimental situations that only allow the accurate measurement of reflection magnitude and not phase. As an example, the approach is demonstrated on one-dimensional numerical simulations that simulate uncertainty in target location and ambient noise. [Work supported by ONR.]

2:00

3pSA3. Backscattering by tilted cylindrical shells above the coincidence frequency caused by helical flexural waves. Florian J. Blonigen and Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814)

Generalized leaky Lamb waves excited by ultrasound incident on cylindrical shells in water produce important backscattering effects. At a certain critical tilt angle, leaky waves launched along the meridian of the cylinder have been shown to lead to large backscattering enhancements [S. F. Morse *et al.*, J. Acoust. Soc. Am. **103**, 785–794 (1998)]. At angles less than the critical tilt, these leaky waves follow helical paths and also make significant contributions. In the present work, experiments using tone bursts were performed on a stainless steel shell to investigate the contributions of a_0 leaky Lamb modes. The tone bursts were of sufficient duration to superpose helical wave bursts of successive circumnavigations, along with the meridional contribution near the critical tilt, to arrive at a steady state backscattering amplitude for the cylinder. This was compared against an approximate numerical partial wave series solution and a ray theory solution as a function of the tilt angle. The data follow the basic shape of the theoretical curves. However, the experimental amplitudes are

more suppressed at higher frequencies resulting from losses due to reflection of leaky waves at the cylinder's end. [Work supported by the Office of Naval Research.]

2:15

3pSA4. Acoustic scattering of metallic springs in elastic medium. Jinmin Zhang (DSO Natl. Labs., 20 Science Park Dr., Singapore 118230), Vijay K. Varadan, and Vasundara V. Varadan (Penn State Univ., State College, PA 16801)

In applications such as ultrasonic immersion testing, damping of flow noise in pipes and ducts and coating of acoustic control surfaces, it is very useful to reduce the reflection of acoustic waves from surfaces. Significant echo reduction can be achieved by addition of an anechoic coating. According to some general design guidelines for a single layer of homogenous coating, a single layer of 15 dB coating must be at least one wavelength thick. For absorbing acoustic waves of low frequencies, inhomogeneities such as metallic fillers, cavities and irregular metallic helices can be introduced into the lossy matrix to substantially enhance the acoustic absorption by acoustic scattering. In this presentation, the acoustic scattering of a single metallic spring embedded in an infinite elastic medium has been investigated using a boundary element formulation. The cross sections of the scattered longitudinal and shear waves have been calculated for various spring sizes, orientations and wave numbers. It will be shown that helical inclusions are more efficient in scattering longitudinal incident waves than spherical inclusions of equal volume. Steel helices are found to be more efficient in converting the incident longitudinal waves into shear waves than helical cavities of equal size.

2:30

3pSA5. Lamb waves scattering by inhomogeneities in elastic waveguides. Natalya S. Gorodetskaya and Victor T. Grinchenko (Inst. of Hydromechanics NAS of Ukraine, 8/4 Zhelabov St., Kiev 252057, Ukraine, vin-igm@gu.kiev.ua)

Different types of conjunctions of elements with various mechanical properties are widely used in engineering practice. Such conjunctions strongly affected the process of acoustic energy transmission in constructions. To study the influence of the physical and geometrical discontinuities in conjunctions on the wave propagation process the composite waveguides have been considered. The method of the boundary problem solution taking into account local singularities of stress is presented. The quantitative data for characteristics of transmitted and reflected waves at the conjunction are given in a wide band of frequencies. At relatively low frequencies the estimations of energy flows by beam model are correct. When the frequency increases, the wave field near conjunction becomes more complicated. As a result the ratio between transmitted and reflected energy becomes a sharply changing function of frequency even in the domain when only one propagating wave exists. This phenomenon is caused by strong excitation of evanescent waves. Excitation of these waves was due to strong mismatch between the form of the incident wave and first transmitted and reflected waves. At the frequency when transmitted energy was minimum the mismatch was maximum. At the same frequency the maximum of normal stresses on the conjunction has been found.

WEDNESDAY AFTERNOON, 6 DECEMBER 2000 NEWPORT ROOM NORTH, 1:00 TO 3:00 P.M.

Session 3pSC

Speech Communication and Signal Processing in Acoustics: Signal Processing for Speech (Poster Session)

Shrikanth S. Narayanan, Chair

AT&T Laboratories-Research, 180 Park Avenue, Room D106, Florham Park, New Jersey 07932

Contributed Papers

All posters will be on display from 1:00 p.m. to 3:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:00 p.m. to 2:00 p.m. and contributors of even-numbered papers will be at their posters from 2:00 p.m. to 3:00 p.m.

3pSC1. Intelligibility of aircraft loudspeakers. Glenn E. Warnaka, Mark E. Warnaka (Future Technologies, LLC 1612 S. Allen St., State College, PA 16801), Peter Mapp (Peter Mapp Assoc., Colchester C03 4JZ, UK), and Bruce Shimizu (C&D Aerosp., 5412 Argosy Dr., Huntington Beach, CA 92649)

Commercial aircraft audio systems provide audible messages including greetings, details of the flight, comments on entertainment, and important safety information. Hence, the speech intelligibility of aircraft audio systems is of great importance. Distributed-mode loudspeakers (DMLs) are known to have wider sound dispersion, especially at higher frequencies, than conventional cone loudspeakers. As aircraft loudspeakers are located relatively close to the passengers, the narrower dispersion and beaming of the high frequencies from conventional cone loudspeakers can cause loss of intelligibility for passengers not on the axis of the loudspeaker and produce uneven sound distribution within the cabin. DMLs have demonstrated intelligibility superior to conventional loudspeakers in other applications such as public address systems. DMLs have other advantages in aircraft installations. They are very lightweight and can even be built into the internal trim, resulting in a nearly zero-weight loudspeaker. Their essentially flat configuration also requires less volume. This paper presents a comparative evaluation of the speech intelligibility of current aircraft loudspeakers and DMLs as determined by the STI (Speech Transmission Index) using the RASTI (Rapid Speech Transmission Index) method in accordance with RTCA document DO-214. Tests were performed under representative conditions and for different seating arrangements.

3pSC2. Improved speech intelligibility in noise with a single-microphone noise reduction technique. Michael J. Nilsson, Xiaoling Fang, Robert M. Ghent, Jr., Patrick Murphy, and Victor H. Bray , Jr. (SONIC innovations, 2795 E. Cottonwood Pkwy., Ste. 660, Salt Lake City, UT 84121)

Reception threshold for sentences (RTS) as measured by a modified version of the HINT test [Nilsson *et al.*, J. Acoust. Soc. Am. **95**, 1085–1099 (1984)] were collected on 26 hearing-impaired listeners fit binaurally with digital hearing aids incorporating a nine-channel spectral subtraction technique of single-microphone noise reduction. Thresholds were measured in quiet and noise presented at a zero degree azimuth with the

subjects listening unaided, aided without noise reduction, and aided with noise reduction and no modification to the gain function when noise reduction was activated. Additional data included soundfield thresholds as a frequency-dependent measure of the impact of the noise reduction algorithm on sensitivity to soft sounds. RTS in quiet reveal a significant effect of noise reduction [$F(3,69)=46.9$, $p<0.01$] with noise reduction increasing RTS 3.4 dB relative to no noise reduction, from 40.4 dB (A) to 43.8 dB (A). RTS in noise reveal a significant effect of noise reduction [$F(3,69)=18.7$, $p<0.01$] with noise reduction decreasing RTS 1.8 dB relative to no noise reduction, from 0.0 dB S/N to 1.8 dB S/N. Soundfield thresholds reveal a loss of sensitivity mainly in frequencies below 1000 Hz. The noise reduction algorithm will be reviewed and discussed.

3pSC3. Automatic speech recognition for mobile hand-held devices. Richard Rose, Shrikanth Narayanan, S. Parthasarathy, Aaron Rosenberg (AT&T Labs-Res., 180 Park Ave., Florham Park, NJ 07932), and Bojana Gajic (Norwegian Univ. of Sci. and Technol., Trondheim, Norway)

The implementation and evaluation of an automatic speech recognition (ASR) based task for a mobile, hand-held device is presented. The use of ASR on mobile devices will make environmental noise, and the design of transducers and compensation algorithms for dealing with noise, critical issues for the success of ASR based services on these devices. A description of the task is provided along with a set of compensation techniques that are used to compensate speaker independent hidden Markov models (HMMs) with respect to environment and transducer variability. Data were collected in a prototype application environment for a form-filling directory access task with speech and pen input and text output (48 subjects in an office environment; 21 in a cafeteria). A technique for combined environment/transducer compensation is presented for this task and is shown to significantly reduce the effects of environmental mismatch. The overall performance degradation was reduced from 41.7% to 10.4% for speech spoken through a far-field microphone in an office environment, and from 79.2% to 39.8% for the same transducer in a noisy cafeteria environment.

3pSC4. Soft-decision speech signal estimation. Leonid Krasny (Ericsson, Inc., Development Dr. 7001, Research Triangle Park, NC 27709, euslekr@rtp.ericsson.se)

Speech is often corrupted by background noise which degrades the performance of coding and recognition algorithms. This scenario is especially true in hands-free mobile communications where noise can be generated by the car engine, the road, and wind hitting the windows. It is essential to reduce the noise level without distorting the original speech signal. A well-known technique for solving this problem is spectral subtraction (SS), where the power spectrum of the noise is estimated when speech is absent and used to suppress noise when speech is present. Normally, a voice activity detector (VAD) is used to determine the presence of speech. In the conventional approach, VAD and SS algorithm act independently. Specifically, the speech signal is first detected by the VAD, and once the presence of the speech is detected, the SS algorithm is used to estimate the signal. Although each task can be optimized separately, the resulting algorithm is not optimal. In this paper, we present a novel approach for estimating the speech signal in which the detection and SS tasks are combined optimally. Our approach employs a soft-decision algorithm for detecting and estimating the speech signal, which leads to improved speech quality.

3pSC5. Time-dependent signal representations that are independent of sensor calibration. David N. Levin (Dept. of Radiol., Univ. of Chicago, Chicago, IL 60637, d-levin@uchicago.edu)

Suppose that the sensor arrays of two speech recognition systems, Sx and Sy, are calibrated differently so that they have different responses (x and y) to the same sound. It is difficult for pattern recognition software to handle such variable sensory data. This paper shows how to avoid this problem: each device transforms its sensory data elicited by an evolving sound into the *same* ''canonical'' representation so that it is suitable for analysis by pattern recognition algorithms. Assume that the two systems have previously been exposed to samples of the same collection of sounds. These ''training'' data induce an affine-connected differential geometry on the sound manifold, which the two sensor arrays represent in different coordinate systems (x and y). Without communicating, the two systems can utilize this geometry to construct the same coordinate-independent (scalar) representation of any time-dependent sound of interest. This methodology is demonstrated by applying it to simulated data from devices with two sensory channels (e.g., for measuring the frequency and amplitude of tones). This approach mimics human perception, which also tends to be constant across systems (individuals) with different sensory distortions, as long as they have been exposed to similar collections of stimuli in the past.

3pSC6. Crime-detection speaker recognition system. Igor I. Gorban (Inst. of Math. Machines and Systems, 42 Ave. Acad. Glushkov, Kiev, 03187 Ukraine, gorban@immsp.kiev.ua), Vitalij N. Magera (Police State Res. Ctr. of Crime-Detection Expertise, Kiev, 01024 Ukraine), and Sergey V. Levij (Natl. Tech. Univ. of Ukraine, Kiev, 04055 Ukraine)

Crime-detection automatic speaker verification and identification (CASVI) system oriented to nonprofessional experts was described at the 137th ASA Meeting [I. I. Gorban, N. I. Gorban, and A. V. Klimenko, J. Acoust. Soc. Am. **105**, 1353 (1999)]. At the current meeting another speaker recognition system for professional crime-detection experts is presented. It consist of special units, Pentium-type computer, and programs realized speech-processing algorithms in automated and automatic modes. The system gives the possibility to research records made on different tape recorders, to fit tape speed, to decode speech records, to suppress noises, to correct frequency and nonline record distortions, to sort sounds, to measure and calculate different discriminate features, characteristics, and parameters usually used by crime-detection experts in their practical work, to compare the records, to form libraries and archives, and to make reports. The system has a simple and comfortable interface tailor-made for crime-detection expert aims. Police exploiting of the system has shown that utilization of the system raises productivity of expert work sometimes and essentially improves authenticity of phonoscope expertise.

3pSC7. Thai monophthongs classification using CDHMM. Ekkarit Maneenoi, Somchai Jitapunkul, Visarut Ahkuputra (Digital Signal Processing Res. Lab., Dept. of Elec. Eng., Faculty of Eng., Chulalongkorn Univ.), and Sudaporn Luksaneeyanawin (Chulalongkorn Univ.)

This paper presents Thai monophthongal vowels recognition. The Thai monophthongs were qualitatively recognized by the 3-state left-to-right continuous density hidden Markov model. The 18 monophthongs are qualitatively 9 different vowels, each of which has two members, short and long. The LPC cepstral coefficients were used and the temporal cepstral derivative was additionally utilized to compare efficiency of the additional feature with the single feature. Qualitative recognition means that short and long vowel pairs were categorized in the same model. Thai polysyllabic words were used in this research. The database consists of 2100 training phonemes from 30 speakers and 1378 testing phonemes from a different group of 20 speakers, respectively. The highest recognition rate of the single feature obtained from 18-order LPC cepstral coefficients is 86.983 percent, while the recognition rate of the 16-order LPC cepstral coefficients plus temporal derivative is 94.580 percent. The results indicate that all the LPC cepstral coefficients associated with temporal derivative have better recognition accuracy than those of LPC cepstral coefficients. It is concluded that the additional temporal derivative can improve recognition rate. The misclassification is analyzed and it is concluded that this resulted from excessively overlapped distributions of vowels in the low- and back vowel groups, respectively.

3pSC8. A machine learning technique for evaluating cues to stop place. Madelaine C. Plauche' (Dept. of Linguist., 1203 Dwinelle Hall, Univ. of California, Berkeley, CA 94720)

This paper is situated in a long line of phonetic studies that seek to determine and qualify the acoustic cues humans use to identify stop place. Acoustic information relevant to the identification of stop place resides in the relative amplitude (Cooper *et al.,* 1952; Ohde *et al.,* 1983) and spectral characteristics of the burst (Blumstein *et al.,* 1979; Stevens, 1999), as well as in formant transitions (Dorman *et al.,* 1996) and voice onset time (Klatt, 1975). Consonant confusion studies suggest that vocalic context is an important factor in the perception of neighboring stops. In the environment of a high front vowel (/i/), for example, cues such as formant transitions and VOT may be neutralized, resulting in a dependence on transient features that are secondary cues in other contexts (Plauche' *et al.,* 1997). The present study estimates the effect of vocalic context on the relative ranking of acoustic cues for stops in CV sequences using a machine learning algorithm (decision trees) that evaluates a database of 1500 (English) CV tokens and their values for the acoustic features: (1) VOT, (2) energy and power of the burst and release, (3) spectrum at the burst, and (4) formant transitions into the following vowel.

3pSC9. Pitch determination by fundamental-harmonic extraction using MUSIC algorithm. Takahiro Murakami and Yoshihisa Ishida (Meiji Univ., 1-1-1, Higashi-mita, Tama-ku, Kawasaki, 214-8571 Japan)

Short-term analysis pitch determination was executed by using multiple signal classification algorithm, which is an eigen-based subspace decomposition method proposed by Schmidt (Ph.D. thesis, Stanford University, 1981). The MUSIC spectrum, based on subspace principles, is sharply peaked at the frequencies of the sinusoidal components of speech signals without almost receiving the influence of the noise. Since the harmonic structure of the power spectrum of speech signals becomes unclear in a high-frequency domain, fundamental-harmonic extraction is performed using the band-limited MUSIC spectrum. The influence of the noise decreases further by limiting the frequency domain to be analyzed, and calculation time has been shortened greatly. The IDFT of the logarithmic MUSIC power spectrum exhibits a strong peak at the position equal to the pitch period like the cepstral method. Results of pitch determination for male and female Japanese vowels illustrate that the proposed method is more excellent than the cepstral method and can estimate the pitch frequency not being influenced by the noise.

3pSC10. Measurement and visualization of speech quality with VIPER. Harald Mundt and Peter Daniel (Neutrik Cortex Instruments, Erzb.-Buchberger-Allee 14, D-93051 Regensburg, Germany, daniel@neutrik-cortex.de)

In this presentation the new VIPER speech quality measurement (VSQM) will be presented. Audible differences between two versions of a sound can be visually inspected in VIPER through the help of auditory distance spectrograms. The objective instrumental measures provided by VIPER will be validated with the results of listening experiments on speech quality (MOS) and compared with the performance of other speech quality measures (PAMS, TOSQA, MNB, PSQM+, PSQM). The comparison is based on the ITU-T Series P. supplement 23 database. The database has been developed in order to characterize the performance of the ITU-T speech codec G.729. Therefore, a wide range of different speech codes, in noisy and clean conditions, in various tandem conditions, with bit and frame errors of various intensity, was tested for different languages (English, French, Japanese). In order to quantify the assessment performances, the linear correlation coefficient and the standard deviation based on the mapped objective data and the subjective data are used.

3pSC11. Comparison between physical and auditory parametrization of speech corpus for the unit selection in text-to-speech synthesis. Minoru Tsuzaki (ATR Spoken Lang. Translation Res. Labs., 2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-0288 Japan, minoru.tsuzaki@slt.atr.co.jp)

It is important to design an appropriate cost function to improve the quality of speech produced by a corpus-based text-to-speech (TTS) synthesis. Although the final product of the TTS system is evaluated perceptually, the definition of cost functions has to be based on the physical parameters of speech signal. And a cost function will become inappropriate if the discrepancy between the physical and the perceptual measures becomes critically large. The discrepancy could be caused by two main factors, i.e., (1) inappropriate selection of parameters, and (2) inappropriate scaling. One may judge that the selection of physical parameters is inappropriate if changes in a single physical parameter result in changes in multiple perceptual attributes. One may also judge that the scaling of physical dimension is inappropriate if the mapping function of the physical space to the perceptual space appears to be critically nonlinear. Aiming at a better definition of cost functions, a signal analysis based on the standard linear prediction coding (LPC) method and one based on the models of the peripheral auditory system (DSAM-AIM) are compared. The results of principal component analysis for Japanese speech databases of a corpus-based TTS system, CHATR, are presented.

WEDNESDAY AFTERNOON, 6 DECEMBER 2000 SCHOONER/SLOOP ROOMS, 1:10 TO 3:00 P.M.

Session 3pUW

Underwater Acoustics: Propagation

Michael G. Brown, Chair
University of Miami, RSMAS-AMP, 4600 Rickenbacker Causeway, Miami, Florida 33149

Chair's Introduction—1:10

1:15

3pUW1. Sound propagation using impulsive sources on the New Jersey Continental Shelf. Mohsen Badiey (Univ. of Delaware, Newark, DE 19716) and William M. Carey (Boston Univ., Boston, MA 02215)

An acoustic experiment was performed on the New Jersey Continental Shelf in a well surveyed area with known bottom and sub-bottom properties. The measurement system was a calibrated vertical array 14–18 km from a known geological feature. Calibrated airgun shots (AG) and special small omni-directional explosives shots (SO) were used to generate impulsive sounds monitored by a calibrated source hydrophone. The repeatability of the SO and AG shots was excellent. The sea state during this experiment was between 4 and 6 on the Beaufort scale and the ship was held in position on either side of the geological feature. The sound velocity profiles were stable over the measurement period due to the fully developed seas. A mixed layer depth of approximately 20–30 m was followed by a steep thermocline to a depth of 48 m with a gradient of approximately 0.5 sec and a relatively isothermal layer to the bottom which ranges between 70 m at the source location and 80 m at the receiving array. Measurements received on the vertical array were variable but over short periods of time remarkably consistent. Time of flight enabled the separation of the first three modes whose shape and arrival times were to agree with calculations. Comparisons of the arrival structure changes due to the feature and estimates of the backscatter from this feature are presented in relation to the mode arrival statistics for different groups of shots.

1:30

3pUW2. Measurement-based geoacoustic modeling for the Acoustic Characterization Test III. Ilya Rozenfeld, William Siegmann (Dept. of Mathematical Sci., Rensselaer Polytechnic Inst., Troy, NY 12180, rozeni@rpi.edu), William Carey (Boston Univ., Boston, MA 02215), and Peter Cable (GTE/BBN Technologies, New London, CT 06320)

The Acoustic Characterization Test III was performed in the oceanographically complex Strait of Korea to provide accurate measurements of sound transmission and coherence (array signal gain) under known environmental conditions. Bottom sampling and sub-bottom surveys, coupled with archival geophysical information, provided the basis for geoacoustic profiles of sound speed, density, and attenuation versus depth. These profiles and the measured bathymetry and water sound speed were used as input parameters for parabolic equations computations. Very good agreement was obtained between measured and calculated narrow-band transmission loss by employing slightly modified geoacoustic profiles and an attenuation profile with a near-water-sediment interface power-law frequency dependence of 1.8. This power law was determined through the use of an effective attenuation coefficient and a least-square analysis of transmissions from five narrow-band tones between 47 and 604 Hz. These results are consistent with measurements in other sandy–silty areas. Using these parameters, comparisons were made between independent measurements and calculations of broadband transmissions and of signal spread, and excellent agreement was found. [Work supported by ONR.]

1:45

3pUW3. Focused arrivals in shallow water propagation. Harry A. DeFerrari, Neil J. Williams, and Hien B. Nguyen (Univ. of Miami, Rosenstiel School of Marine and Atmospheric Sci., 4600 Rickenbacker Cswy., Miami, FL 33149)

Observations of acoustic pulse transmissions for several sites in the Florida Straits consistently show that a single intense pulse dominates the arrival pattern. For these sound channels, the propagation is by refracted modes that interact with the bottom below the critical angle resulting in favorable low-loss transmission. The intense arrivals come about from focusing of the waveguide and are always observed for downward refracting sound speed profiles whenever the source and receiver are deep in the channel. The foci correspond to caustics of ray theory or to the phase-stationary summation of many modes. Such arrivals are often 20 dB higher than others and are of interest in that they will be the only detectable features at long ranges and have been shown to be stable and persistent enough for tomographic inversion. The coherence properties and fluctuation statistics of focused arrivals are presented and analyzed. Data from a 28-day propagation experiment are presented. The receptions of broadband pulse transmissions are compared for a 10-km propagation range for pulses with center frequencies from 100 to 3200 Hz. Simultaneous measures of the sound speed field are used as model inputs to predict and compare spatial and temporal features of the focused arrivals.

2:00

3pUW4. Anomalous signal loss in the Yellow Sea, revisited: Coupling the acoustics with model-generated oceanographic realizations. Stanley A. Chin-Bing, David B. King, Alex C. Warn-Varnas, Robert A. Zingarelli (Naval Res. Lab., Stennis Space Center, MS 39529-5004, chinbing@nrlssc.navy.mil), and Kevin G. Lamb (Univ. of Waterloo, Waterloo, ON N2L 3G1, Canada)

In a seminal paper, Zhou *et al.* [J. Acoust. Soc. Am. **90**, 2042–2054 (1991)] introduced the concept of large anomalous signal loss due to a resonance effect caused by solitons (internal waves traveling along the thermocline). They assumed that solitons were present and produced a remarkable comparison between acoustic model predictions and their acoustic measurements taken over a four-year period. They did not take sufficient oceanographic measurements to confirm the existence of solitons. Numerous investigators have tried to duplicate their pioneering work in similar shallow-water environments with solitons present. Signal reduction, mode conversions, and resonance-like effects have been observed, but signal loss of the magnitude measured by Zhou *et al.* have not been observed. We have used a primitive equation soliton model and the tidal flow near the Shandong peninsula to generate soliton simulations that flow into the Yellow Sea near the region where Zhou *et al.* made their acoustic measurements. We are performing analysis similar to that of Zhou *et al.* on these soliton realizations to determine if large signal losses can occur. Results from this investigation will be presented and discussed. [Work supported by ONR/NRL and by a High Performance Computing DoD grant.]

2:15

3pUW5. Broadband acoustic variability due to internal solitary waves. Yongke Mu, Mohsen Badiey (College of Marine Studies, Univ. of Delaware, Newark, DE 19716), James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), Stephen N. Wolf (Naval Res. Lab., Washington, DC 20375), and William L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180)

A shallow water (70–90-meter-deep) broadband acoustic experiment with a source-receiver range of 15–18 km in an azimuthally dependent environment was designed as a part of the SWARM'95 field study. Temporal behavior of the water column was sampled every minute from two sources placed above and below the thermocline transmitting signals over a period of several hours while the water column was measured for the passage of internal waves. Spatial behavior was sampled in the vertical plane by hydrophones with spacing on the order of meters. Coherence of broadband acoustic waves for frequencies (20–300 Hz) is examined for the waveguide over different environmental conditions. The effect of environmental variability on coherence, in particular the sound-speed fluctuations in the water column due to the internal solitary waves, is noted as a function of acoustic frequency and azimuth. Analysis of the acoustic fluctuations over short time scales (10–15 minutes) may resolve the temporal decorrelation of the received signal due to internal waves. The vertical sampling of the received signal permits an analysis of arrival-angle fluctuations. Numerical simulations are performed, which provide support to the experimental observations. [Work supported by ONR.]

2:30

3pUW6. Ray dynamics in the AET experiment. Michael G. Brown (Rosenstiel School of Marine and Atmospheric Sci., Univ. of Miami, Miami, FL 33149, mbrown@rsmas.miami.edu), John A. Colosi (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), Anatoly L. Virovlyansky, George M. Zaslavsky (Courant Inst. of Mathematical Sci., New York Univ., New York, NY 10012), Steven Tomsovic, and Michael A. Wolfson (Washington State Univ., Pullman, WA 99164)

A ray-based wave field description is employed in the analysis of measurements made during the November 1994 Acoustic Engineering Test (the AET experiment). In this experiment phase-coded pulse-like signals with 75-Hz center frequency and 37.5-Hz bandwidth were transmitted near the sound channel axis in the eastern North Pacific Ocean. The resulting acoustic signals were recorded on a moored vertical receiving array at a range of 3252 km. In our analysis both mesoscale and internal-wave-induced sound speed perturbations are taken into account. Much of this analysis exploits results that relate to the subject of ray chaos; these results follow from the Hamiltonian structure of the ray equations. It is argued that all of the important features of the measured AET wave fields are consistent with a ray-based wave field description in which ray trajectories are predominantly chaotic. [Work supported by ONR.]

2:45

3pUW7. Time variability of acoustic signals in a benign shallow-water area. Peter L. Nielsen, Martin Siderius, and Finn B. Jensen (SACLANT Undersea Res. Ctr., Viale San Bartolomeo 400, 19138 La Spezia, Italy)

SACLANTCEN conducted the ADVENT'99 experiment in the Strait of Sicily, Mediterranean, to assess the time variability of broadband acoustic signals in a benign shallow-water region. Broadband acoustic signals (200–3800 Hz) were transmitted for up to 18 h over fixed paths of 2, 5, and 10 km. Dense sampling of the environment was performed including a sound-speed section every hour along the acoustic tracks. Band-averaged transmission loss is stable with time, which agrees well with modeling results. The Bartlett processor was applied to correlate the acoustic data. The correlation time varies from several hours at low frequency and short range to a few minutes at high frequency and longer ranges. Range- and time-dependent propagation modeling of the acoustic signals (200–800 Hz) shows behavior similar to the data. Particularly, the correlation time decreases abruptly from several hours to less than 1 hour at a frequency around 700 Hz for the 10-km data. This effect of the time-varying waveguide on the acoustic signals is correctly predicted by the propagation model. Severe problems using the acoustic signals coherently persist above a few hundred Hz and at ranges beyond a few kilometers, although the experiment was conducted under favorable conditions. The experimental data and the simulations illustrate the complexity of sound propagation in shallow water.

Plenary Session and Awards Ceremony

Katherine S. Harris, President
Acoustical Society of America

Presentation of Certificates to New Fellows

R. Dean Ayers
Soren Bech
Peter D'Antonio
Laurent J. D. Demany
Mark A. Holden
Donald E. Hall
Gordon R. Hamilton
W. Jack Hughes
Dennis F. Jones
Richard H. Love
Charles Thompson
Ji-Qing Wang
Oleg A. Godin
Hari S. Paul
Hans C. Strifors

Presentation of Awards

Science Writing Award in Acoustics for Journalists to Kathryn Brown

Science Writing Award in Acoustics for Journalists to Roland Pease and Radek Boschetty

Science Writing Award for Professionals in Acoustics to William M. Hartmann

Distinguished Service Citation to John V. Bouyoucos

Distinguished Service Citation to F. Avril Brenig

Pioneers of Underwater Acoustics Medal to Darrell R. Jackson

Silver Medal in Physical Acoustics to Gregory W. Swift

Session 4aAA

Architectural Acoustics: Theme Park Acoustics

David E. Marsh, Chair
Pelton Marsh Kinsella, 1420 West Mockingbird Lane, #400, Dallas, Texas 75247-4932

Chair's Introduction—8:25

Invited Papers

8:30

4aAA1. Challenges and more in the acoustic design of an attraction. Neil A. Shaw and Ray Van den Broeck (Menlo Sci. Acoust., Inc., P.O. Box 1610, Topanga, CA 90290-1610)

Building a theme park from scratch involves not only providing a ''theme'' for the attractions contained therein, but also integrating the acoustic control and enhancement measures so that they work with, and not against, the goal of the attraction, which is the suspension of disbelief, at least for the age group to which the attraction is designed. The challenges encountered by the authors, the analysis of the same, as well as the proposed solutions, in the acoustic design for an enclosed attraction that includes hands-on interactive activities, live performances, and a roller coaster will be reviewed. Field work to determine potential mitigation paths will also be discussed.

9:00

4aAA2. Creativity in theme park acoustics. Marshall Long (Marshall Long Acoustics, 13636 Riverside Dr., Sherman Oaks, CA 91423, m_long@pacificnet.net)

Theme parks are places where the entire environment is a theater in which the guests are active participants. As such it emphasizes creativity of concept and design, including acoustical design. Creative solutions have been found for acoustical problems including innovative uses of materials, mechanical and sound systems. A number of examples are presented from the author's 20-year experience with theme park acoustics illustrating examples of creative use of scientific principles.

9:30

4aAA3. Knott's Camp Snoopy at the Mall of America: A theme park under glass, revisited. Steven J. Thorburn (2867 Grove Way, Castro Valley, CA 94546, SJT@TA-Inc.com)

Over ten years ago, the design of a special theme park was born. The plans for Knott's Camp Snoopy at the Mall of America in Minneapolis, Minnesota would eventually grow into a popular tourist destination, but there were initial concerns about acoustics. The owners were familiar with the high noise levels of other indoor amusement parks and wanted to avoid that same situation. Having just completed the West Edmonton Mall, where many complaints were received about noise levels and guest visits were very short, it was very important to the owners that the design would prevent problems associated with noise. To do this, an acoustical consulting engineering firm was brought in. In this paper, the original acoustical design goals and solutions will be reviewed and compared against the final constructed project and its success over the first decade.

10:00

4aAA4. Community noise from theme parks. Ted N. Carnes (Pelton Marsh Kinsella, 1420 W. Mockingbird Ln., Ste. 400, Dallas, TX 75247)

Theme and amusement parks are made to generate fun and excitement. As part of the guest experience a great amount of sound energy is often generated and it may be perceived as noise in the surrounding community. Defining outdoor noise criteria for the parks is often done for you by the local municipality zoning and building officials. Thus, one can use sound-modeling techniques to determine potential problems before the parks are built or expanded. While noise control of sound levels at the source is always desirable, there are many sources that defy their use. Thus, control of sound levels along the sound path is all that is often available. All of these concerns about theme parks will be addressed along with some suggestions on how to minimize community noise impacts. Some actual project examples will be discussed with the architectural acoustics aspects emphasized.

10:30

4aAA5. The impact of propagation anomalies on noise emissions from entertainment facilities. Robert Bronsdon (Walt Disney Imagineering, 1401 Flower St., Glendale, CA 91221)

Most measurements made on environmental noise sources are done whenever it is convenient without any real consideration to the effects of environmental factors such as wind and temperature gradients. Predictions are typically done assuming that the atmospheric conditions are stable and benign. When the predictions and test results are compared and do not match, a generalization is given: ''It

must have been caused by some type of atmospheric condition.'' A computational Gaussian beam program that accounts for the effects of wind and temperature inversions was presented at Inter-noise 98 in New Zealand and is now currently a part of the SoundPLAN program for evaluating environmental noise. That program is designed primarily to deal with temperature inversions and wind that is blowing from the source to the receiver, but it does provide insight into upwind propagation. This paper explains how upwind propagation of sound in temperature inversion conditions can impact noise receivers in urban settings.

11:00

4aAA6. Noise abatement strategies and modeling approaches for outdoor attractions at theme parks. Christopher W. Menge (Harris Miller Miller & Hanson, Inc., 15 New England Executive Park, Burlington, MA 01803, cmenge@hmmh.com)

Recent experience in noise abatement approaches and results for a new theme park are presented. Significant contributions from different source components on roller coasters lead to different abatement strategies taken for each. Measured source data will be presented. The relative performance of abatement applied separately to the source and sound path will be discussed, as well as the effects of source directivity. Source and propagation modeling approaches are presented, including how a customized version of a recently developed national noise model for highway traffic was applied for roller coasters and other sources.

THURSDAY MORNING, 7 DECEMBER 2000 SCHOONER/SLOOP ROOMS, 7:25 A.M. TO 12:05 P.M.

Session 4aAB

Animal Bioacoustics: Instrumentation for Animal Bioacoustics Monitoring and Measurement

William C. Burgess, Chair
Greeneridge Sciences, 1411 Firestone Road, Goleta, California 93117

Chair's Introduction—7:25

Invited Papers

7:30

4aAB1. Reflector microphones for field recording of natural sounds. George W. Swenson, Jr. (Dept. of Elec. and Computer Eng., Univ. of Illinois at Urbana-Champaign, 1308 W. Main St., Urbana, IL 61801, gswenson@uiuc.edu)

A directional microphone system for field recording of sounds in the air, for example, bird song or other animal activity, usually involves either a parabolic reflector to focus the sound waves on the microphone (transducer) element, or a linear array of transducers so phased as to respond preferentially to sounds from one directional sector. The latter system (the so-called ''shotgun'' microphone) can be analyzed in a fairly straightforward manner. The reflector system, however, involving as it does a structure comparable to a wavelength in linear dimension, is not susceptible to the conventional approximate methods of computation. Recently developed computational techniques now permit exact calculation of the directional responses of small reflectors. One result is a proposal for a very economical and effective system involving a plane reflector. No directional microphone can, in practice, reproduce sounds with fidelity to the sounds as emitted by the source.

7:50

4aAB2. Portable instrumentation for the recording, analysis, and playback of infrasonic animal vocalizations. Elizabeth von Muggenthaler (Fauna Commun. Res. Inst., P.O. Box 1126, Hillsborough, NC 27278)

Several bioacoustic experiments required an affordable system capable of field recording, real-time analysis, signal editing, and real-time playback of infrasonic animal vocalizations. The synthesis of several types of recording devices and signal processing software has resulted in the capacity to record infrasound and perform analysis even during rain, heat, and high humidity. Small portable microphones and recorders can record in the field from 3 Hz to 22 kHz. On-site analysis including real-time FFT, color spectrographic function, filtering, cross-correlation, and other functions can be facilitated with the use of any portable computer with 92 Mbytes RAM. Signal editing, including frequency, amplitude, and cut/paste can be accomplished with readily available music signal processing software. Real-time field playback from 10 Hz to 45 kHz can be acheived by using portable car audio speakers. This system was used for the real-time analysis of elephant vocalizations, which were edited and played back immediately during recording sessions. It has also been used in field research involving tiger, binturong, rhinoceros, and giraffe. Tigers and elephants appear to respond behaviorally to real-time playback of original and edited vocalizations. [Work supported by National Instruments, Momentum Data Systems, and Sonic Foundry.]

8:10

4aAB3. Animal sound analysis and playback software: Building a software tool useful to field researchers. John M. Burt (Cornell Lab of Ornithology, 159 Sapsucker Woods Rd., Ithaca, NY 14850)

Many researchers would like to conduct sound-playback experiments with the level of control potentially available using a computer, and yet there are few programs specifically designed for computer-based playback. The lack of PC playback software led to the development of a WINDOWS 98 application called SYRINX, which is specifically designed for both sound analysis and playback in the field. The goal was to provide a host of playback features that would allow researchers to conduct sophisticated field playback experiments using a sound card-equipped laptop PC and a portable speaker. SYRINX can display multiple sound files as playable spectrograms, has a small suite of tools for editing sounds, displays a real-time scrolling spectrogram of audio input, allows users to log behavioral events, and has a number of other features useful for playback research. The talk will focus on the field-research-driven design goals for the SYRINX project, how the program was implemented, what signal-processing tools and features were thought to be necessary for field researchers, what new features may be implemented in the future, and the challenges involved in developing a researcher- and field-friendly playback program.

8:30

4aAB4. Geographically distributed acoustical monitoring of migrating birds. Harold Mills (Bioacoustics Res. Prog., Cornell Lab. of Ornithology, 159 Sapsucker Woods Rd., Ithaca, NY 14850, hgm1@cornell.edu)

A geographically distributed system of largely automated acoustical monitoring stations was developed to monitor the migration of small passerine birds. The targeted species, wood warblers and sparrows, migrate at night and produce short (about 120 ms at the longest), high-pitched (between roughly 6 and 10 kHz) calls as they fly, often audible from the ground. The monitoring system consisted of ten stations, located in New Jersey, Maryland, Pennsylvania, and New York, and a central data repository located at the Cornell Lab of Ornithology. Each station comprised an outdoor, weatherproof microphone connected to a volunteer's home computer. The computer automatically ran acoustic transient detection software each night that listened continuously to the microphone signal and extracted all transients matching certain criteria, each to its own audio file. The detected transients were uploaded by the volunteer each morning via the Internet to the central repository for classification and archival. Results, including spectrograms of all detected transients and bird call counts, were displayed on a public web site updated daily. [Work supported by EPA.]

8:50

4aAB5. A compact digital recording system for autonomous bioacoustic monitoring. Thomas A. Calupca, Kurt M. Fristrup, and Christopher W. Clark (Bioacoustics Res. Prog., Cornell Lab. of Ornithology, 159 Sapsucker Woods Rd., Ithaca, NY 14850)

A digital recording system with several months' endurance was designed to monitor biological sounds. The system combined off-the-shelf components (microcontrollers, notebook hard drives) with an IDE interface designed for this purpose. The present storage capacity of the system is 25 Gbytes, with a maximum sampling rate of 50 kHz at a resolution of 12 bits. Power consumption during typical long-term continuous recording averages 250–500 milliwatts, and its physical dimensions are approximately $3\times7\times10$ cm, including the hard drive. Hardware and software support for GPS clock synchronization enables correction of clock drift in the system, as well as precise synchronization of multiple units in an array. This unit has been extensively utilized in bottom-moored recording buoys, called ''popups.'' Popups include an acoustic communication system that receives and acknowledges commands from a surface support vessel, and an electronically controlled galvanic release. Everything is housed in a single 17-in. glass pressure housing. More recently, these systems have been used to monitor low-frequency terrestrial signals in studies of forest and savannah elephant populations. An upcoming project will drift these systems over inaccessible habitats using helium balloons.

9:10

4aAB6. An autonomous acoustic recorder using a directional sensor for locating calling bowhead whales. Robert G. Norman and Charles R. Greene, Jr. (Greeneridge Sci., Inc., 1411 Firestone Rd., Goleta, CA 93117)

A new directional acoustic recorder has been developed to assess the movement and distribution of vocalizing marine mammals. Called DASAR (Directional Autonomous Seafloor Acoustic Recorder), the instrument incorporates an omnidirectional and two orthogonal horizontal sensors from a DIFAR (DIrectional Frequency and Recording) sonobuoy, as well as a magnetic compass. With appropriate processing, the horizontal directions can be computed for received sounds. The acoustic bandwidth is 1000 Hz, appropriate for bowhead calls, and recording is continuous for 45 days. DASARs were deployed in a hexagonal array seaward of an artificial gravel island recently constructed for oil and gas production northwest of Prudhoe Bay, Alaska. Operation was during the fall migration of bowhead whales past the island. The objective was to relate calling-whale locations to the character and levels of sounds emanating from the island and from nearby tugs and barges. [Work supported by BP Exploration (Alaska).]

9:30

4aAB7. A digital acoustic recording tag for measuring the response of marine mammals to sound. Mark Johnson, Peter Tyack, Douglas Nowacek (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, majohnson@whoi.edu), and Alex Shorter (Univ. of Colorado, Boulder, CO 80302)

There is growing concern about the effect on marine mammals of underwater sound sources such as vessels, industrial equipment, and sonar. In order to quantify the reaction of these animals to sound, controlled exposure experiments have been attempted using surface observations or vocal monitoring to estimate response. However, the short surface time of most marine mammals, and their often unknown vocalization rates, limit the effectiveness of such experiments. To address these problems, a small digital recording tag, the DTAG, has been developed capable of simultaneously sampling the acoustic environment of the host animal, together with its orientation, heading, and depth. The tag has a 400-Mbyte memory array sufficient to record audio and sensor signals for several hours.

The tag is encapsulated in a plastic resin and can operate at a water depth up to 3000 m. Programing and data off-load are accomplished with an infrared communications link to a personal computer. The DTAG has been deployed on northern right whales and sperm whales using suction cup attachments. Delivery was by means of a cantilevered pole from a small boat. The resulting data indicate a number of behavioral response metrics as well as new results on dive behavior and vocalization rates.

9:50

4aAB8. The bioacoustic probe: A general-purpose acoustic recording tag. William C. Burgess (Greeneridge Sci., Inc.)

Recent research has demonstrated the utility of acoustic recording tags to assess the exposure and response of marine wildlife to sound [e.g., W. C. Burgess *et al.*, Deep-Sea Res., Part II **45**, 1327–1351 (1998)]. The experimental nature of these tags, however, has limited their adoption by the wider bioacoustical community. To enable broader use of acoustic recording tag technology, a new general-purpose tag is being developed. Miniaturization to a cylinder approximately 3 cm in diameter by 16 cm in length will allow application with a variety of species and attachment methods. Initial versions of the device will sample acoustics with 16-bit resolution at bandwidths up to 14 kHz, as well as temperature and depth with 12-bit resolution. Longevity will depend on the choice of sampling schedule; constant acoustic sampling at 2 kHz will fill the 288 MB solid-state flash disk in 21 h, but this lifetime can be extended by reducing resolution or by recording only during times of interest. Results from initial test deployments on free-ranging marine mammals are expected in late fall 2000. [Work supported by ONR.]

10:10–10:20 Break

Contributed Papers

10:20

4aAB9. Echolocation signals of wild white beaked dolphins measured with a four-hydrophone short base line array in real-time. Whitlow W. L. Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734), Marianne H. Rasmussen, and Lee Miller (Ctr. for Sound Commun., Odense Univ., Odense, Denmark)

The echolocation signals of Lagenorhynchus albirostris were measured in the waters off Keflavik, Icelend during 22–26 August 1998. A four-hydrophone array was used with each hydrophone located at the tip of an equilateral triangle and one located at the center of the triangle. Each of the outer three hydrophones was spaced 0.61 m from the center hydrophone and the separation angle between outer hydrophones was 120 deg. An underwater video camera was attached 10 cm from the center hydrophone. The array was mounted on a long pole that was lowered to a depth of about 2 m from the side of a 10-m diesel driven boat. The hydrophones were connected to two GAGA-1210 data acquisition boards operating at a 500-kHz sample rate. The computer clock was synchronized with the clock on portable video recorder. Two-hundred pretrigger and 200 post-trigger points per channel with four channels of data were acquired simultaneously and stored in the memory of a transportable computer. The results indicate that the white beaked dolphins emitted high-frequency clicks with peak frequencies close to 120 kHz. The maximum source level of the echolocation signals was 219 dB and the source levels increased with range from the array.

10:35

4aAB10. Miniature wildlife telemetry transmitter equipped with a microphone. Larry Pater (USA CERL, 2902 Farber Dr., Champaign, IL 61821), William Cochran (Univ. of Illinois), and George Swenson, Jr. (Univ. of Illinois)

A miniature microphone was integrated into a telemetry transmitter package with a total weight of 1.2 grams (operating life of 7 days) or 2.0 grams (21-day life). To obtain meaningful acoustical information, the radio transmitter is continuous transmission, rather than the pulsed type typically used for wildlife telemetry. It is sensitive enough to pick up bird songs at a distance of over 30 m and small airplane engine sounds at several kilometers. This device is useful for remote monitoring in inaccessible locations and to avoid the intrusive impact of direct observation by humans. It can be used to document acoustical environment and auditory stimuli and to monitor vocalizations and behavior of the animal on which the transmitter package is mounted. Applications include studying behavior and responses to disturbance. The transmitter has been successfully used to monitor bird calls and wing-beat patterns during various types of activity, including foraging and nighttime migration flights.

10:50

4aAB11. Use of a radio microphone array to study banded-wren song interactions at the neighborhood level. John M. Burt (Cornell Lab. of Ornithology, 159 Sapsucker Woods Rd., Ithaca, NY 14850)

An array of distributed microphones was used to locate and identify singing banded wrens in Santa Rosa National Park, Costa Rica. Wrens were recorded both during natural singing interactions and during song playback experiments. The array consisted of 11 weatherproofed radio-microphones distributed among neighborhoods of up to seven banded-wren territories, with a centrally located radio receiver/recording station. Audio from the radio receivers was stored to a laptop computer using a multichannel data-acquisition card and custom recording software. Singing birds were located after recording using software developed at the Cornell Lab of Ornithology's Bioacoustics Research Program. The system's durability and flexibility enabled several different array configurations to be tested at two field sites under a range of recording conditions. Preliminary results suggest that this system is a powerful and flexible tool for studying the dynamics of vocal interactions across entire neighborhoods of singing birds.

11:05

4aAB12. Remote infrared acoustically triggered video monitoring unit. Robert Ambrose (U.S. Fish and Wildlife Service, 1011 E. Tudor Rd., Anchorage, AK 99503), Michael Donaldson (Far North Aquatics, Box 61264, Fairbanks, AK 99706), Timothy Lavallee, and Troy Andersen (Geo-Marine, Inc., Newport News, VA 23606)

A remotely operable, infrared system for recording digital video and audio information immediately before, during, and after an acoustical event was developed. The primary application of this system to date has been the recording and documentation of effects of noise on wildlife, including Peregrine falcons (*Falco peregrinus*), Mexican spotted owls (*Strix occidentalis lucida*), and brown pelicans (*Pelecanus occidentalis*). The system triggers, records, and stores digital video and audio data on an IBM compatible computer running WIN9X. Serial communication with a Larson Davis sound-level meter allows a field video camera to store video and audible sound data on the computer. The need for field crew collection of bioacoustical effects data at sites with video acquisition systems is minimized. The video information reduces observation or interpretation bias in the data collection procedures. The systems can be maintained from remote locations using cellular modem technology and can be operated without being visited for extended periods. It is low powered and can be operated by a system of batteries and solar panels for indefinite field deployment. With the correct lens the system is capable of seeing in complete darkness up to 25 ft or as far away as 200 ft in the daytime.

11:20

4aAB13. Programmable DSP datalogger for animal bioacoustics monitoring. David A. Mann and Timothy J. Tucker (Tucker-Davis Technologies, 4637 NW 6th St., Gainesville, FL 32609)

Two hardware modules for animal bioacoustics monitoring based on the Analog Devices Sharc DSP are described. The RP2 Real-Time Processor connects to a PC via the USB port and can be used for signal detection and direct logging to disk. The SD2 Smart Datalogger is a battery-powered digital datalogger. The RP2 has two channels of A/D with 100-kHz bandwidth, two channels of D/A with 200-kHz bandwidth, eight digital inputs, eight digital outputs, and up to 64 MB RAM. The SD2 Smart Datalogger has two channels of A/D and D/A with a 25-kHz bandwidth on each channel, two digital I/O channels, and up to 64 MB RAM. The datalogger weighs 2 ounces and its dimensions are $5.6\times7.6\times1.8$ cm^3. DSP programming for both devices is accomplished through the use of a drag-and-drop interface for configuring detection circuits using a library of over 120 DSP components including digital filters, signal detectors, and memory operations. The use of the RP2 for detecting and recording animal sounds in the field will be described where sounds in given frequency bands are detected, time-stamped, and saved to memory or hard disk. [Work supported by NIH.]

11:35

4aAB14. The development of a compact instrument for the measurement of biological sea noise. Alec J. Duncan (Ctr. for Marine Sci. and Technol., Curtin Univ. of Technol., Bentley, Western Australia), Douglas H. Cato (Defence Sci. and Technol. Organisation, Pyrmont, NSW 2009, Australia), Frank Thomas, and Robert D. McCauley (Curtin Univ. of Technol., Bentley, Western Australia)

This paper describes a compact, low-power instrument that has been developed to monitor biological sound sources and other sea noise over extended periods in remote locations. The instrument is based on a low power microcontroller and was designed to provide flexibility in specifying sampling rates and sampling schedules so that the utilization of battery power and disk space could be optimized for any given application. In its basic configuration the signals from two hydrophones are sampled and stored on a 12-GByte hard disk drive. Multiple drives can be installed if greater capacity is required. The electronics fits inside a 100-mm i.d. underwater housing resulting in a package that is easy to transport and deploy. Further details of the instrument design will be given, together with the results of deployments off the Western Australian coast where it has been used to monitor the passage of migrating humback whales.

11:50

4aAB15. Detection of harbor porpoises and white whales (Beluga) sound using the high frequency bioacoustics and their echo structures. Erhan Mutlu (Inst. of Marine Sci., METU, P.O. Box 28 Erdemli, 33731 Turkey, mutlu@ims.metu.edu.tr)

Acquisition of sounds emitted by marine mammals (porpoise and beluga) living in the Black Sea was made using active echosounder (BioSonics Model 102) with down-looking 120- and 200-kHz transducers. The sounds of the mammals displayed different structures not similar to each other. The individual amplitude of the porpoise sound was significantly discriminated and varied between -55 and -62 dB at 200 kHz. The echoes were characterized with decreasing amplitude from the deep water to the surface. It looked like vertical lineations. The sound scattering was fluctuated in a range of 10–20 dB. Sound emitted by beluga displayed a structure like rakes of a comb. The echo was captured first very weak at 120 kHz and then strong at 200 kHz. The volume scattering changed within a range of 60 and -70 dB at 200 kHz. While the whale was close to the transducer, it emitted the sound every 300 s. It sometimes gave sound in different patterns from some reason. As we got away from the whale, while on the way, we detected the sound at 10 s. Each lineation of the comb-rake structure appeared in a certain time delay and the strength decreased toward the surface.

THURSDAY MORNING, 7 DECEMBER 2000 CALIFORNIA SALON 3, 7:50 TO 11:30 A.M.

Session 4aAO

Acoustical Oceanography: Acoustical Oceanography and Satellite Remote Sensing

Daniel L. Hutt, Chair

Defence Research Establishment Atlantic, P.O. Box 1012, Dartmouth, Nova Scotia B2Y 3Z7, Canada

Chair's Introduction—7:50

Invited Papers

8:00

4aAO1. Estimation of subsurface ocean density structure using remote sensing and data assimilation. Josko Bobanovic and Keith Thompson (Dept. of Oceanogr., Dalhousie Univ., Halifax, NS B3H 4J1, Canada, josko@phys.ocean.dal.ca)

Recently, researchers from Dalhousie University, The Royal Military College, and the Defence Research Establishment Atlantic started work on estimating water density, and eventually underwater acoustical properties, from *in situ* and remotely sensed data. This requires the development of assimilation schemes to project vertically into the ocean interior the surface information gathered by remote sensing. Initially, we are concentrating on inferring the seasonal distributions of temperature and salinity on the continental shelf. We are experimenting with new ways of assimilating such data into fully nonlinear, baroclinic models using the incremental approach. In a nutshell, the approach uses two parallel ocean models: a complex model to be fit to the data, and a simple model with known adjoint that is used to correct the control variables of the complex model. The result is a set of seasonal fields that is consistent with the observed data, and dynamical constraints imposed by the model, within prescribed errors bars. Our long-term objective is to develop practical schemes for assimilating time sequences of remotely sensed (e.g., SST, ocean color, altimetry, HF radar) data into any ocean model. Progress will be reviewed.

8:30

4aAO2. Using MODAS to derive subsurface sound speed estimates from satellite observations. Charlie N. Barron, Daniel N. Fox (Naval Res. Lab., Oceanogr. Div., Stennis Space Center, MS 39529), and Eleanor Holmes (IPD/A&T/Anteon, Arlington, VA 22202)

The Modular Ocean Data Assimilation System (MODAS) provides a set of tools for analysis and manipulation of ocean data. One of its functions is the estimation of subsurface temperature and salinity fields from satellite observations of height deviation and temperature. A two-dimensional optimal interpolation of the satellite data produces fields of surface height anomaly and temperature. These data are provided to the MODAS synthetic profile routine to predict subsurface temperature and salinity fields compatible with the surface observations and local climatological statistics. Available *in situ* measurements may be assimilated at this stage. A sound speed field calculated from the temperature and salinity estimates provides the basis for additional acoustical applications. MODAS products, evaluated by comparison with finely sampled surveys, show significant improvement over estimates from static climatologies. [This work was funded by the Space and Naval Warfare Systems Command (SPAWAR) and the ONR through the Advanced Data Fusion Center Support and the On-Scene Tactical Ocean Forecast Capability projects (SPAWAR, Program Element 603207N) and the MODAS Improvements and the Prediction of Coastal Buoyancy Jets projects (ONR, Program Element 602435N).]

9:00

4aAO3. MEANS: Multiscale environmental assessment network studies. Henrik Schmidt (MIT, 77 Massachusetts Ave., Cambridge, MA 02139), Allan Robinson (Harvard Univ., Cambridge, MA 02139), Lakshmi Kantha (Univ. of Colorado, Boulder, CO), and Jurgen Sellschopp (SACLANT Undersea Res. Ctr., La Spezia, Italy)

The small-scale coastal acoustic environment is extremely difficult to predict and forecast. The very shallow near-shore region, with water depths of up to a few tens of meters, is particularly challenging due to the variability of the critical seabed properties due to sediment transport processes, biological processes, and changes in small-scale bathymetry due to wave action, with strong implications for the high-frequency reverberation characteristics in particular. The potential of using small autonomous underwater vehicles (AUV) with high dependency on acoustic systems in such environments has led to an increased need for forecasting the acoustic characteristics. The coastal ocean environment is a result of a very complex dynamical system with strong coupling among different scales. To investigate the predictability of this process, the MEANS effort implements a nested forecasting framework for the small-scale oceanography in Procchio Bay, Elba during the GOATS'2000 AUV technology experiment in Sep.–Oct. 2000. A local high-resolution model is constrained by assimilating oceanographic data from AUVs and moorings, and nested within a regional Ligurian Sea model and a basin scale circulation model, both constrained by surface ship measurements and satellite remote sensing data. The nested observation and modeling framework will be described, and preliminary results will be presented and discussed. [Work supported by ONR and SACLANT.]

9:30

4aAO4. The dependence of low-frequency underwater surface scattering on remotely sensed oceanographic variables. Roger C. Gauss (Naval Res. Lab., Washington, DC 20375-5350, roger.gauss@nrl.navy.mil), Eric I. Thorsos, Frank S. Henyey (Univ. of Washington, Seattle, WA 98105-6698), and Joseph M. Fialkowski (Planning Systems, Inc., McLean, VA 22102-3304)

Recent low-frequency (<1500 Hz) underwater acoustic measurements have demonstrated that scattering processes at the air–sea interface depend primarily on the wind speed and the surface wave spectrum. The scattering strength from the rough interface is proportional to the spectral density at the Bragg wavelength with modifications due to tilt and modulation by longer waves. These modifications are accounted for in the small slope approximation. When wave breaking becomes significant, rough interface scattering is augmented by bubble cloud scattering, which depends primarily on wind speed. In this regime, bubble cloud scattering dominates at low grazing angles, and rough interface scattering dominates at high grazing angles. A physics-based empirical model is used to describe bubble scattering. The mean-frequency-shift characteristics of acoustic signals scattered from both the moving sea surface and bubble clouds have been successfully modeled given the 2-D surface wave spectrum. These scattering-strength and frequency-shift models are used to explore the sensitivity of low-frequency scattering to environmental variables obtainable by remote sensing. [Work supported by ONR.]

10:00

4aAO5. The estimation of acoustic properties of internal waves using synthetic aperture radar (SAR). Justin Small and Mark Prior (Defence Evaluation Res. Agency, Winfrith Technol. Ctr., DT2 8XJ, UK)

This paper looks at some methods relevant to the assessment of the acoustic properties of internal waves from satellite data. The currents in these internal waves strain the sea surface leading to signatures, which can be imaged by instruments such as radar, photography, and SAR. Two lines of approach for estimating internal wave amplitudes from SAR will be illustrated. The first approach uses length-scale information on SAR (separation of waves, width of waves), combined with some knowledge of the stratification (from a database or survey) and a solitary internal wave or undular bore model. The second approach uses radar imaging models to investigate the strength of current which would lead to the observed signatures in the given wind conditions. One important acoustic property of internal waves is the influence on surface duct propagation due to the large measured excursions of the mixed layer. Examples will be shown of the effects on surface duct propagation for different frequencies and scenarios using experimental data taken (1) off the northwest of the UK; (2) in the central Bay of Biscay; and (3) on the Portuguese coast.

10:30–10:45 Break

10:45

4aAO6. Horizontal refraction modal tomography (HRMT) for ocean front and ocean current monitoring. E. C. Shang and Y. Y. Wang (CIRES, Univ. of Colorado/NOAA/Environ. Technol. Lab., 325 Broadway, Boulder, CO 80303)

It has been pointed out by Voronovich and Shang [IEEE J. Ocean. Eng. **24**, 224–231 (1999)], that the HRMT has the ability of reducing the mode-coupling effect. In this paper, two examples of the application of HRMT in strong mode-coupling environment are discussed: (1) Polar front, and (2) transverse current in Fram Strait. In both cases, strong mode coupling takes place for the used frequency and mode number. Due to the modal phase difference is measured by two close paths, the mode-coupling effect can be significantly reduced. The results of the retrieved frontal parameters and the transverse component of the current based on adiabatic theory are pretty good even without any iteration. For the Polar front, the horizontal refraction angle is estimated. The parameters that can be retrieved by using HRMT are: the normal direction of the front, the center location of the front, and the sound speed profile difference of the two water masses. The interaction of those acoustically estimated parameters with satellite remote sensing is discussed. [Work supported by NOAA and ONR.]

11:00

4aAO7. A comparison of acoustic and TOPEX/POSEIDON altimetric data in the Kuroshio Extension region during summer 1997. Gang Yuan, Humio Mitsudera (IPRC-SOEST, Univ. of Hawaii, 2525 Correa Rd., Honolulu, HI 96822-2219), Iwao Nakano, Hidetoshi Fujimori, Yasushi Yoshikawa, and Toshiaki Nakamura (Japan Marine Sci. and Tech. Ctr., Yokosuka, 237-0061 Japan)

Approximately two months of data from a five-transceiver acoustic tomography array in the Kuroshio Extension (KE) in the summer of 1997 are used for comparing with the TOPEX/POSEIDON (T/P) altimeter measurements. The acoustic data show that travel time (TT) variations are very consistent with contemporaneous variations of sea surface height anomalies (SSHA) as in the T/P in the KE front region, but have significant differences in the KE recirculation region. This difference is assumed to be due to the fact that the altimeter might not pick up the signal associated with the variations occurring below the seasonal thermocline. The inversions from a new approach show that the temperature in the surface layer (0–100 m) had been warming up 1.0 °C during the experimental period, but cooling down about 0.2 °C in the subsurface layer (100–1500 m). This temperature warming up in the surface layer derived from acoustic data is excellent, consistent with the steric height change determined from net heat flux data of the NCEP/NCAR. Zonal range and depth averaged (barotropic) westward-flowing current velocity determined from reciprocal acoustic TTs in the recirculation region is about 5 cm/s; it is comparable with the surface geostrophic current velocity determined from the SSHA of the T/P. Acoustic tomography complements the altimetry.

11:15

4aAO8. Application of space-based synthetic aperture radar imagery for estimating underwater ambient noise. Daniel L. Hutt (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada, daniel.hutt@drea.dnd.ca)

Synthetic aperture radar (SAR) imagery of the ocean surface provided by the Radarsat satellite can provide estimates of surface waves, winds, and shipping density. These are the parameters which are mainly responsible for underwater acoustical ambient noise. The remotely sensed surface conditions can be combined with knowledge of the acoustical propagation environment and geoacoustic properties of the bottom to estimate the ambient noise field. Although difficulties exist with SAR-derived wind fields and ship detection, these are balanced by the wide area coverage and high spatial resolution of the SAR sensor. We review space-based SAR capabilities relevant to ambient noise estimation and compare Radarsat-derived wind fields to *in situ* measurements for two Radarsat passes over the Atlantic Ocean. Ambient noise levels measured during the overpasses compare favorably with those estimated using the Radarsat wind data.

THURSDAY MORNING, 7 DECEMBER 2000 PACIFIC SALONS A, B AND C, 7:55 TO 11:00 A.M.

Session 4aEA

Engineering Acoustics: Acoustic Noise Characterization and Mitigation

James O. Buntin, Chair
Brown–Buntin Associates, Inc., 7996 California Avenue, Suite A, Fair Oaks, California 95628

Chair's Introduction—7:55

Contributed Papers

8:00

4aEA1. Gas turbine noise mitigation for a residential development. Jim Buntin (Brown–Buntin Associates, Inc., 7996 California Ave., Ste. A, Fair Oaks, CA 95628)

A residential development was proposed adjacent to a gas turbine electrical power production peaking facility. To determine compliance with local standards, noise levels and frequency content were measured at reference locations under summer and winter conditions. The Environ. Noise Model (ENM) was calibrated on the basis of the reference noise level data, and noise measurement results at other locations were compared to the ENM predictions. The measurement results were also used to determine the effective noise source height for calculations of shielding. The ENM was then used to predict noise levels at the development site. To evaluate methods of obtaining compliance with the standards, the ENM was used to assess the effects of berms, walls, and grading on the development site, and an enclosure for the gas turbines. Recommendations were then prepared for noise abatement and mitigation measures, including the use of an enclosure, an onsite berm, and reduction of low frequency noise.

8:15

4aEA2. Computational investigation of wave steepening in a two-dimensional duct. W. R. Erskine, A. Selamet, and P. M. Radavich (Dept. of Mech. Eng., The Ohio State Univ., Columbus, OH 43210)

The need for accurate modeling of nonlinear phenomena, such as wave steepening, in unsteady flows is important in many engineering applications. Wave steepening causes the initial sinusoidal shape of the oscillation to distort and eventually form a shock wave. Both pressure and velocity waves steepen as flow moves downstream inside a duct. A computational fluid dynamics code, which solves the compressible, unsteady, two-dimensional Navier–Stokes equations, is used to investigate this nonlinear behavior and predict the flow field in a long narrow duct. A sinusoidal inlet velocity and an anechoic termination outlet are imposed as boundary conditions. The shape and frequency content of the traveling wave are predicted at various locations inside the duct. Inlet velocity is varied to cover a range of small-to-large amplitude pressure fluctuations. Results indicate that wave steepening is enhanced as the wave travels downstream and/or as the fluctuation amplitude increases. At large enough amplitudes and/or duct lengths, a shock is produced which travels to the outlet. Regardless of the oscillation frequency of the inlet velocity, the frequency spectra reveal sound-pressure level amplitude peaks at multiples of both quarter- and half-wavelength resonances based on the length of the duct.

8:30

4aEA3. Fast acoustic power computations for structural optimization problems. John B. Fahnline (Applied Research Lab., 16 Appl. Sci., University Park, PA 16802)

Boundary element analyses in acoustics are often computationally inefficient because a large complex-valued matrix has to be solved at every frequency for which a solution is desired. This problem can be alleviated somewhat by using appropriate mesh densities for both the structural acoustic analyses, as discussed previously by the author. To achieve further enhancements, a computer program has been written to interpolate the acoustic resistance matrix over a frequency band from values calculated at sparse intervals. A preliminary analysis is performed to map out the resistance matrix as a function of frequency, requiring specification of only the frequency band of interest and the desired error tolerance. Automatic refinement is performed in regions where the resistance varies rapidly, as it does near acoustic resonances. A simple interpolation scheme is then used to compute the resistance matrix at any arbitrary frequency knowing the precomputed matrices, resulting in an order of magnitude reduction in the computation times when compared to a direct calculation. Because the acoustic analysis is independent of structural parameters, this formulation is ideal for problems where structural parameters are to be optimized without changing the radiating surface's shape. Examples are given to demonstrate the formulation's accuracy and computational efficiency.

8:45

4aEA4. Lagging application with use of sound intensity. Sokcheat Monh (Brown–Buntin Associates, Inc., 7996 California Ave., Ste. A, Fair Oaks, CA 95628)

Improvements were proposed to an existing county waste water sump station. The improvements included installation of a scrubber unit used to filter odor from the sump chambers. Noise from the operation of the scrubber unit raised concerns among nearby residents. To address compliance with local standards and minimize impacts on nearby homes, sound intensity testing was used to determine which modules of the scrubber unit made significant contribution to the overall noise levels. Lagging was recommended, and it was implemented by the county. Before and after sound intensity levels of the scrubber unit demonstrated the effectiveness of lagging as a mitigation measure.

9:00

4aEA5. Computation of free-jet turbulence and exhaust-mixing noise from a turbofan engine. David B. Schein (Northrop Grumman Corp., 9L53/W6, One Hornet Way, El Segundo, CA 90245-2804 and Dept. of Mech. and Aerosp. Eng., UCLA, scheida@mail.northgrum.com) and William C. Meecham (UCLA, Los Angeles, CA)

A model presented previously [D. B. Schein, paper 3pEA3, 136th Meeting of ASA (1998)] has been extended and applied to prediction of exhaust noise from a high Reynolds number, hot exhaust flow. Simulations for a Williams Research WR19-4 turbofan engine exhaust have been performed and results compared with far-field acoustic data measured during static test runs of the engine at exhaust Mach numbers up to 0.8. The hybrid CFD/CAA approach employs a deductive subgrid scale model for small-scale compressible turbulence combined with the time-derivative form of Lighthill's source-integral acoustic analogy, containing the fluctuations represented by the time-varying stress tensor. Overall sound pressure level directivity and representative spectra are presented. Analytical results indicate generally good agreement with measured data and popular semiempirical prediction methods. Spectral analysis indicates that significantly larger computational domains and simulation times will be required to obtain low-frequency energy content radiated from aircraft engine exhausts.

9:15

4aEA6. Noise and vibration control for the Alameda Corridor project. Joel Vazquez (Parsons Eng. Sci., Inc., 100 W. Walnut St., Pasadena, CA 91124, jvazquez@parsons.com)

The Alameda Corridor project is a 20-mile railroad cargo expressline that will connect the ports of Los Angeles and Long Beach to the transcontinental rail network east of Los Angeles. Once completed, the corridor will create a fast and efficient way to move cargo through United States and overseas world markets. The midcorridor section of the project consists of a 10-mile long, 33-ft descent into a by-railway Trench. Although this Trench will be the featured portion of the midcorridor project, the construction of a by-pass track will facilitate cargo-laden freight passage while the Trench is constructed as well as provide access for local businesses after completed. On its path from the ports to downtown Los Angeles this massive construction project will traverse potential noise and vibration-sensitive areas that house single and multi-family residences, motels, schools, daycare centers, parks, and churches. One of the greatest challenges encountered with projects of this magnitude is the successful control of construction and operational noise and vibration. The presentation will focus on various noise and vibration control methods that have been implemented as part of an abatement program. The scope of this successful program includes monitoring, equipment noise certification, and impact studies.

9:30–9:45 Break

9:45

4aEA7. The multiple source feedback of supersonic jet screech, I: Theory—frequency. Alan Powell (Dept. of Mech. Eng., Univ. of Houston, Houston, TX 77204-4792)

Screech sound, generated by the interaction of jet instability waves with the spatially periodic cell structure of choked or ''off-design'' supersonic jets, feeds back to initiate new instability waves at the nozzle. The interaction produces a phased array of *effective* point sources distance s apart, with a single equivalent (centroidal) source at distance $n's$ from the nozzle. The feedback criteria discussed at the last Meeting yield frequency $f=f_{PR}\phi$, where $f_{PR}=(U/s)\cdot 1/(1+M)$ for *perfect reinforcement* at the nozzle; $U,M=$instability wave velocity, Mach number; $\phi=(N+p+\nu+I)/n'\rightarrow 1$ with $n's\equiv(N+\mu+\nu)s$, $\nu s=$distance of equivalent source from nearest cell end at $(N+\mu)s$ from the nozzle, $N=$integer, $|\mu|\leqslant 1/2$,

and $I=-1,\ 0$, or $+1$. Exact for any number of symmetrically disposed source strengths, numerical solutions show it to be a good approximation for arbitrary strengths and for inclusion of differing source distances (with weighted equivalent source). The mean s is appropriate for unequal source spacing; but then perfect reinforcement cannot occur. With perfect reinforcement unlikely, the 1953 hypothesis that $f \rightarrow f_{PR}$ is confirmed as a simple and useful approximation.

10:00

4aEA8. The multiple source feedback of supersonic jet screech, II: Space-time diagrams and the standing wave puzzle elucidated. Alan Powell (Dept. of Mech. Eng., Univ. of Houston, Houston, TX 77204-4792)

Space-time diagrams for defined phases of the periodic instability wave and the feedback sound waves from the resultant *equivalent* point sources provide a simple physical picture of the multiple source feedback process, illustrated for cases of perfect reinforcement when $\phi=1$, and for $p=1/2$ and for $\mu=\pm 1/2$ when $\phi \neq 1$. They also provide a simple interpretation of the nature of the (sometimes necessary) approximation. The intersections of the maxima of the instability and sound waves form ''partial standing waves'' with maxima $L_{SW}=s/\phi$ apart, involving one complete cycle of the instability and sound wave. This provides a simple explanation of the puzzling standing waves observed by Westley and Woolley (1975), Rice and Taghavi (1992), and by Panda (1996, 1997), it having been speculated that they may indicate, or be fundamentally associated with, the actual sound sources. Rather, $n's \equiv (N+\mu+\nu)s=(N+p+\nu)L_{SW}$ showing, for $\nu=0$, that $(N+p)$ standing waves (cycles) must fit in the nozzle-to-source distance, closely analogous to the N'th stage of (high-speed) edge tones. Then both μ and p can be estimated from such experimental data.

10:15

4aEA9. The multiple source feedback of supersonic jet screech, III: Space-time diagrams applied to the helical mode. Alan Powell (Dept. of Mech. Eng., Univ. of Houston, Houston, TX 77204-4792)

The foregoing feedback theory applies directly to the sinuous and toroidal instability modes of supersonic jets, but needs modification for the helical C mode since the acoustic feedback to the nozzle must at the axial trace velocity of the resultant helical sound field (Powell, Umeda, and Ishii, 1990). Here the space-time diagram is equivalent to an instantaneous map of the positions of the helical instability and acoustic wave on the cylindrical bounding surface of the jet, one period being replaced by one circumference. The instantaneous space-time (sound) pressure measurements of Westley and Woolley (1975) are interpreted in this way. The acoustic feedback trace velocity is found to be approximately sonic. The circumferential velocity ($=\pi f \times$diameter source orbit diameter), in a plane normal to the jet axis is inferred to be supersonic. The phase velocity of the helical wave along the curved surface is then found to be subsonic rather than sonic as originally assumed in the tentative theory. A simple numerical simulation (including the reflection from the flange) reproduces the main features of the complex map of instantaneous pressure, including the prominent standing waves.

10:30

4aEA10. Hydrodynamic lift and drag fluctuations of a sphere—theoretical model. M. S. Howe (College of Eng., Boston Univ., 110 Cummington St., Boston, MA 02215), G. C. Lauchle (Penn State Univ., State College, PA 16804), and J. Wang (The Trane Co., LaCrosse, WI 54601-7599)

A theoretical investigation is made of the unsteady lift and drag exerted on a sphere in a nominally steady high Reynolds number, incompressible flow. The net force on the sphere has previously been ascribed to fluctuations in the bound vorticity in the meridian plane normal to the force, produced by large-scale coherent structures shed into the wake. A simplified model of vortex shedding is proposed that involves coherent eddies in the form of a succession of randomly oriented vortex rings, interconnected by pairs of oppositely rotating line vortices, and shed at quasiregular intervals with a Strouhal number ~ 0.19. The rings are rapidly dissipated by turbulence diffusion, but it is shown that only the nascent vortex ring makes a significant contribution to the surface force, and that the force spectrum at Strouhal numbers exceeding unity is effectively independent of the shape of the fully formed vortex. Predictions of the lift and drag spectra at these frequencies are found to be in good accord with new towing tank measurements performed on a neutrally buoyant sphere configured as an acoustic velocity hydrophone. [Work supported by ONR Code 321 SS.]

10:45

4aEA11. Hydrodynamic lift and drag fluctuations of a sphere—empirical model. J. Wang (The Trane Co., 3600 Pammel Creek Rd., LaCrosse, WI 54601-7599), G. C. Lauchle (Penn State Univ., State College, PA 16804), and M. S. Howe (Boston Univ., Boston, MA 02215)

An empirical model is developed to estimate the unsteady lift and drag induced on a spherical body subjected to a steady, uniform flow. A wind tunnel is used to measure the statistics of the surface pressure fluctuations on a sphere under subcritical Reynolds number conditions. The measurement results are incorporated into the model for predicting the unsteady force induced on the sphere. The model is based on a separable assumption of the cross-power spectral densities of the surface pressure fluctuations. This assumption is shown to be a proper engineering approximation except in low-frequency ranges. The model functions are obtained using least mean-square curve fits of the data. Tow tank experiments are performed, where the flow-induced unsteady side force and drag are measured independently of each other on spheres configured as acoustic velocity hydrophones. The empirical model and the tow tank measurements are in very good agreement with each other and also with a simplified vortex ring model presented in the accompanying paper. The predicted forces are used to determine the flow-induced noise on underwater inertial sensors subjected to a steady, uniform flow. [Work supported by ONR Code 321 SS.]

Session 4aED

Education in Acoustics and Noise: Public Education in Noise and Acoustics

Daniel R. Raichel, Chair
532 Spencer Drive, Wykoff, New Jersey 07481

Invited Papers

9:00

4aED1. Nonprofits and noise education, the experience of the Noise Pollution Clearinghouse. Leslie D. Blomberg (Noise Pollution Clearinghouse, Box 1137, Montpelier, VT 05601, les@nonoise.org)

During its four-year history, the Noise Pollution Clearinghouse has become a national leader in educating the public about noise and its effects. This paper explains the work of the Noise Pollution Clearinghouse, the reasons for its success, and the unmet needs and opportunities that remain. Particular attention is paid to the potential for nonprofits and professional societies to be leaders in noise education.

9:20

4aED2. Boom cars: Noise pollution at its worst. Daniel R. Raichel (Grad. School, CUNY, New York, NY 10031)

Boom cars, sometimes referred to as ''boomers,'' are vehicles that have stereo sound equipment installed specifically to generate exceedingly loud sound levels. Because the market is so lucrative, particularly among automobile owners under 30 years of age, manufacturers aggressively engage in promotion of ultra-high-power amplifiers, auxiliary bass drivers, midrange drivers, tweeters, crossovers, etc. Loudness competitions are held; and levels as high as 170 dB have been reported. Most advertising emphasizes loudness—for example, a TV spot shows a van's sound system shattering store windows as it speeds down a street. In addition to an extremely high probability of hearing loss, those exposed to the excessive vibration inside a boomer's interior may experience other sorts of physiological trauma, this in addition to annoying the surrounding neighborhood with excruciatingly high sound levels. The phenomenon of boom cars compares to that of tobacco addiction—a danger to vehicle occupants and a nuisance to ''secondary'' listeners. Countermeasures against such public nuisances include: (1) enactment and diligent enforcement of municipal laws specifying limits on discernable sound levels, (2) education in secondary schools on the dangers of excessive noise levels, and (3) public criticism of manufacturers' promotional efforts that emphasize extremely high sound power.

Contributed Papers

9:40

4aED3. Noise control in buildings—A noise-control class for architectural engineers. Ralph T. Muehleisen (Dept. of Civil, Environ., and Architectural Eng., Univ. of Colorado, Boulder, CO 80309)

As public awareness of noise control grows and the public tolerance for noise diminishes, interest in building noise control is increasing. Since noise control is better applied as an integral part of the building design, rather than as an afterthought, building designers need to be educated in the basics of noise control. To this end a new class entitled Noise Control in Buildings has been developed for architectural engineering majors at the University of Colorado (CU). The class is a senior-level/graduate-level design class that introduces the basics of acoustics and noise control as applied to building design. In the class students learn from lecture, homework, guest lecturers from industry, field trips, experimental sessions, and projects. A discussion of the class content and teaching method is provided.

9:55

4aED4. The 1-day public symposium as a format for acoustics dissemination. Sten Ternstrom (Dept. of Speech, Music and Hearing, Kungl. Tekniska Hogskolan, Drottning Kristinas Vag 31, SE-100 44 Stockholm, Sweden, sten@speech.kth.se)

The 1-day public seminar or symposium is a combined education and dissemination format that has been successfully applied by the Music Acoustics group at KTH in Stockholm, about once a year for the past 25 years. Such a seminar is typically structured as a full morning and afternoon of scheduled, invited lectures on a particular theme, followed by a public evening concert or other performance that corresponds to the theme. One or two international guests are keynote speakers, while the bulk of the program is contributed by leading Scandinavian researchers, teachers, and performers. The target style for the presentations is that of an introduction or overview for the interested and motivated but nonexpert academic. The topics are chosen for their expected public appeal, from ongoing or recent research activities in music acoustics and voice acoustics. Recent topics have included ''Music and motion,'' ''Real-time biofeedback in voice therapy and pedagogy,'' and ''Randomness and noise in music.'' Attendance usually ranges from 70 to 200. When possible the lectures are also published in book form, complete with sound examples. Aspects of content, public relations, and the financing of such symposia are discussed.

10:10

4aED5. ''Sound'' science and technical education for sixth graders: A practical approach. Harry Alter (19 Beechtree Ln., Granville, OH 43023) and Rebecca Evans (1655 River Rd., Granville, OH 43023)

Following the renovation of the Granville Middle School Cafeteria in Granville, OH it was noticed by students and faculty that noise levels during lunch period had dramatically increased. As a result, speech became unintelligible and eating became both unenjoyable and uncomfortable. The resident sixth-grade science teacher with the assistance of a local acoustical engineer saw the opportunity to teach the sixth-grade science class about acoustics and to get the children directly involved in solving the school's cafeteria noise problem. The class met on a weekly basis to

learn about sound and its influences on the school environment. Ultimately, the students would build a single microphone impedance tube and test the sound absorption characteristics of various materials. From their investigation they would design and install an acoustical wall treatment system to reduce reverberant noise levels in the school's cafeteria. Acoustical measurements of the cafeteria were taken before and after treatment to compare the student's calculations and final work. The paper also outlines the comprehensive effort involved from securing local grant monies to technical discussions and photographs of the children's work.

THURSDAY MORNING, 7 DECEMBER 2000 YAWL ROOM, 8:15 TO 9:45 A.M.

Session 4aMUa

Musical Acoustics: Modal Analysis Techniques in Musical Acoustics

Daniel A. Russell, Chair
Science and Mathematics Department, Kettering University, 1700 West Third Street, Flint, Michigan 48504-4898

Invited Papers

8:15

4aMUa1. Application of modal analysis and synthesis of reed and pipe to numerical simulations of a clarinet. Matteo Facchinetti (Laboratoire d'Hydrodynamique, Ecole Polytechnique, 91128 Palaiseau Cedex, France, matteo@ladhyx.polytechnique.fr), Xavier Boutillon (Laboratoire d'Acoustique Musicale, CNRS, Universite Paris 6, Paris, France), and Andrei Constantinescu (Laboratoire de Mecanique des Solides, Ecole Polytechnique, 91128 Palaiseau Cedex, France)

A finite-element model of a complete clarinet is proposed in order to simulate the acoustical dynamics of the instrument. Following the work of Pinard and Laine on isolated reeds [unpublished reports of the Ecole Polytechnique], modal analysis of clarinet reeds performed with holographic interferometry has been compared with modal synthesis derived from the FEM. In this linear model, the reed is considered alone first, then coupled to the mouthpiece, and finally to the mouthpiece and the barrel. The good agreement between the two approaches makes the mixed solid-acoustical FEM usable in a numerical simulation of the whole clarinet. The rest of the pipe is modeled by lumped mechanical elements whose parameters are fitted with measurements of the acoustical input impedance. Eigenmodes of the complete system—reed, mouthpiece, barrel, pipe—are computed and damping is subsequently attributed to each mode. The evolution of the system subject to nonlinear boundary conditions—incoming airflow, one-sided contact of the reed against the curved lay of the mouthpiece—is computed in the time domain: at each time step, external forces and internal variables are projected onto the modal basis and the evolution of each modal amplitude is integrated. Findings of the simulations will be discussed.

8:45

4aMUa2. Modal analysis of musical instruments using TV holography. Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115)

Holographic interferometry gives the greatest spatial resolution of any method of modal analysis. However, recording holograms on photographic film is time consuming and often inconvenient. Electronic systems that use CCD TV cameras to record images and form the interferograms in a small digital computer avoid these inconveniences. In a typical electronic TV holography system, a new interferogram is displayed 30 times a second, providing the opportunity to view the interferograms in real time as some parameter is varied. We illustrate this method by describing the vibrational modes of several musical instruments, including bells, gongs, cymbals, drums, violins, cellos, guitars, mandolins, and steelpans.

Contributed Papers

9:15

4aMUa3. Recent work and new axes of research for MODALYS, a sound synthesis program based on modal representation. Cristophe Vergez, Joel Bensoam, Nicolas Misdariis, and Rene Causse (IRCAM, 1 Pl. Igor Stravinsky, 75004 Paris, France)

MODALYS, developed at Ircam after 10 years for research and musical applications, is based on modal representation of vibrating objects, mechanical or acoustical. Modal data (natural frequencies, damping factors, and deflection shapes) are obtained from: (1) analytic modeling for simple structures as ideal string, cylindrical or conical open/closed tubes, rectangular plate with a simply supported boundary condition or violin bridge, . . .); (2) experimental measurements for real structures with complex shapes (violin body, soundboard of piano, flute, . . .); and (3) finally by numerical methods (finite-element methods) which allow the creation of any arbitrary virtual shapes. The vibrating object involved is generally a set of simple and linear substructures. Nonlinearities, when they exist, are introduced in the coupling of substructures, for example reed and tube for the clarinet. Dissipation plays a critical role in the perception of some vibrating musical objects. The type of damping used in our modal representation is referred to as proportional damping (dissipation phenomena are in this case constant or proportional to the square of frequency). However, this type of damping may not apply in any case to the real objects. In this presentation, some sound illustrations on recent results of research and application development with MODALYS will be presented and discussed.

9:30

4aMUa4. Acoustic and modal analysis of an African djembe drum. Daniel A. Russell and Wesley S. Haveman (Sci. & Math Dept., Kettering Univ., 1700 W. Third Ave., Flint, MI 48504)

The African djembe drum consists of a goat skin stretched over a hand-carved shell with a large cavity, open at the bottom. The shape of the shell cavity acts as a Helmholtz resonator, providing a strong bass component around 75 Hz. The shell exibits several bell- or wineglass modes, some of which have frequencies close to membrane modes. This talk will present results from acoustic and modal analysis tests of the drum with and without the skin membrane intact. Mode shapes and frequency spectra will be presented as well as a discussion of how the shell, cavity, and membrane couple together to provide the djembe's unique sound.

THURSDAY MORNING, 7 DECEMBER 2000 YAWL ROOM, 10:00 A.M. TO 12:00 NOON

Session 4aMUb

Musical Acoustics: General Topics in Musical Acoustics

Roger J. Hanson, Chair
Physics Department, University of Northern Iowa, Cedar Falls, Iowa 50614

Contributed Papers

10:00

4aMUb1. Bilingual speakers perceive a musical illusion in accordance with their first language. Diana Deutsch (Dept. of Psych., Univ. of California, San Diego, La Jolla, CA 92093), Trevor Henthorn (Univ. of California, San Diego, La Jolla, CA 92093), and Mark Dolson (E-mu Creative Technol. Ctr., Scotts Valley, CA 95067)

The tritone paradox is a musical illusion that is perceived quite differently by different listeners. Previous findings have indicated that perception of this illusion varies with the language or dialect to which the listener has been exposed, particularly in childhood. In this study, subjects were employed whose first language was Vietnamese and who now reside in California. One group of 16 subjects had arrived in the U.S. as adults. All spoke perfect Vietnamese but little English. Another group of 16 subjects had arrived in the U.S. as children. All spoke perfect English, but most did not speak fluent Vietnamese (though they still heard some Vietnamese spoken). The two groups perceived the tritone paradox in strikingly similar ways, and their perceptions differed significantly from those of another group of Californian subjects who were native speakers of American English. Another experiment, employing fluent Vietnamese speakers, obtained a close correspondence between the way a subject perceived the tritone paradox and the pitch range of his or her speaking voice. The bilingual subjects, therefore, perceived the tritone paradox in accordance with the pitch characteristics of their first language, even when they no longer spoke this language fluently.

10:15

4aMUb2. Laryngeal configuration and acoustical properties during flute vibrato. Joon-Hee Beth Hwang and C. William Thorpe (Faculty of Health Sci., The Univ. of Sydney, East St. Lidcombe, NSW 2141, Australia)

This study reports on simultaneous measurements of glottal aperture and spectral analysis of flute vibrato for a group of six expert flute players. The glottal aperture was recorded using a nasendoscope and the sound was recorded using a portable DAT recorder. Video images were selected from stable portions and digitized, and measurements of the vocal fold opening were made from the display. The sound recording was also digitized and was analyzed using MATLAB. The results showed an oscillatory pattern on the glottal opening together with the vibrato oscillations in the acoustic signal. The change in glottal opening between maximum and minimum adduction during this oscillation was an average of 7.42 mm for loud playing and 1.73 mm for soft playing. The spectral analysis also showed that the amplitude modulation occurs mostly in the higher partials of the tone. These results strongly suggest that the larynx plays an important role in developing vibrato in the flute tone. It appears that the larynx modulates the airflow, thus producing pressure fluctuations in the mouth that generate the oscillatory ''vibration'' of the acoustic tone as observed in previous studies [N. H. Fletcher, J. Acoust. Soc. Am. **57**, 233–237 (1975)].

10:30

4aMUb3. A study of brass instrument ''lipping'' using an artificial mouth. Mark Neal, Francis H. Barnes, D. Murray Campbell, Clive A. Greated (Dept. of Phys. and Astron., Univ. of Edinburgh, Edinburgh EH9 3JZ, Scotland), John S. Cullen (Flat 12, Stow Pk. Ct., Newport NP20 4HB, UK), and Joel Gilbert (Universite du Maine, 72085 Le Mans Cedex 9, France)

Sound is generated in a brass wind instrument as a result of aeroelastic coupling between the mechanical vibratory system formed by the lips and the acoustical vibratory system of the air column. The sounding frequency depends on the parameters of both vibratory systems, and is in general not exactly equal to one of the acoustical resonance frequencies of the air column. A skilled brass instrument player can vary the method of blowing to pull the note above or below an acoustic resonance; this technique is known as ''lipping.'' In order to obtain a clearer understanding of the lipping mechanism, an artificial mouth has been used to drive a trombone mouthpiece coupled to a cylindrical tube of variable length. The phase relationship between mouthpiece pressure and lip opening has been determined for near-threshold auto-oscillation over a range of air column lengths and embouchure adjustments, corresponding to lipping both above and below the acoustic resonance. Implications for the classification of the lip reed mechanism are discussed.

10:45

4aMUb4. Self-sustained oscillation as a function of blowing pressure in a one mass lip reed model. Jeffrey Moffitt and Daniel Ludwigsen (Dept. of Phys., The College of Wooster, 1189 Beall Ave., Wooster, OH 44691, jmoffitt@acs.wooster.edu)

The lip reed is one of the least understood components of the brass player-instrument system. Its biological nature allows a wide range of player control over the parameters that are vital to the regeneration of the pressure wave. The lumped-element model of Adachi and Sato [J. Acoust. Soc. Am. **99**, 1200–1209 (1996)] accomplishes the goal of self-oscillation and displays the correct phase relationships with only two parameters to specify. The embouchure is captured by the natural frequency of the mass-spring system, and the blowing pressure must be sufficient to develop self-sustained oscillation. Now the behavior of a model player-instrument system can be compared with the results of the thorough investigation of blowing pressure of human players by Fletcher and Tarnopolsky [J.

Acoust. Soc. Am. **105**, 874–881 (1999)]. The 2D parameter space is mapped to show regions of self-sustained oscillation as a function of both blowing pressure and the lip's natural frequency. [Work supported by REU Grant NSF-DMR No. 9987850.]

11:00

4aMUb5. Effective lengths of trombone bells with mutes. Michael S. Lodin, R. Dean Ayers, and Lowell J. Eliason (Dept. of Phys. and Astron., California State Univ., Long Beach, 1250 Bellflower Blvd., Long Beach, CA 90840-3901, rdayers@csulb.edu)

A brass player employs mutes in performance to modify spectral characteristics of the sound output for musical effect. An undesirable side effect is a modification in pitch center, which is easily compensated for in the trombone by a small adjustment in playing slide position. The effects of various mutes upon the sounding frequencies of the trombone are studied by determining the change each mute produces in the effective length of an equivalent, closed-end air column. For both the open-bell case as well as with various mutes inserted, an impulse-response technique is used to determine the phase shift of each frequency component induced at the bell end, leading to a calculation of a frequency-dependent effective length for each configuration [R. W. Pyle, Jr., J. Acoust. Soc. Am. **57**, 1309–1317 (1975); R. D. Ayers, *ibid.* **98**, 81–87 (1995)]. The various lengths are compared with each other, as well as with the practical experience of brass players in using them during performance. [Work supported in part by the Scholarly and Creative Activities Committee at CSULB.]

11:15

4aMUb6. Nonlinear effects in an asymmetrical vibrating wire. Roger J. Hanson, H. Kent Macomber, and Andrew Morrison (Dept. of Phys., Univ. of Northern Iowa, Cedar Falls, IA 50614)

The asymmetries of a vibrating harpsichord wire which yield different free vibrational frequencies along two characteristic orthogonal transverse axes were described previously [Hanson *et al.*, J. Acoust Soc. Am. **106**, 2141 (1999)]. This splitting of the fundamental vibrational frequency is of the order of 0.3%. Its value depends upon the particular wire sample and the tension, with lower tensions yielding larger absolute splittings. For the 2nd, 3rd and 4th vibrational modes, splittings differ greatly from integral multiples of the splitting of the fundamental. If the wire is driven sinusoidally there can be large nonlinear generation of motion in a direction (P) perpendicular to the driving direction (D) and of higher harmonics of the driving frequency. These nonlinear effects are, however, very sensitive to the frequency splitting and the relation of the driving direction to the characteristic axes. Theory indicates that $P/D \sim 1$, independent of the vibrational amplitude. However, with a small splitting of 0.11 Hz out of 77 Hz, it was found that to obtain $P/D \sim 1$ an amplitude at the antinode of about 0.15% of the wire length was necessary. The driving frequency region for generation of higher harmonics is sharply dependent on the higher free-vibration mode frequencies.

11:30

4aMUb7. Study on double decay of individual partials of piano sound—Preliminary results on the ''unison quality.'' Rene Causse (IRCAM, 1 Pl. Igor Stravinsky, 75004 Paris, France) and Simon Felix (Universite du Maine, 72017 Le Mans, France)

A characteristic of piano tones is their double decay (the initial sound followed by the ''aftersound''). The difference in mistuning between the three strings of a unison group, and also the adjustment of the regularity of the action of the hammer in the case of important irregularities in hammer striking, are parameters used by the tuner to adjust the temporal evolution of the sound. This study is a part of a more general study to explore what tuners call ''unison quality,'' which means finding objective and qualitative criteria to decide that the temporal evolution is well adjusted. The results of experiments are reported where the decay of the first 12 partials of real sounds recorded under different conditions is analyzed and then modified by signal processing. The results of informal listening tests indicate that experts (tuner, professional pianists, and composers) are sensitive to these modifications (removal of very brief amplitude falls observed for the different partials). The decay of the different partials seems to be more complex than the double decay observed for the global sound and for the fundamental frequency. Sound illustrations will be presented.

11:45

4aMUb8. Nonlinear modeling of the ''mitnahme-effekt'' in coupled organ pipes. Domenico Stanzial, Davide Bonsi, and Diego Gonzalez (Lab. di Acustica Musicale e Architett. FSSG-CNR, Isola di S. Giorgio Maggiore, c/o Fond. Cini, I-30124 Venezia, Italia)

Two slightly mistuned organ pipes of the same nominal tone placed on neighboring holes of the same groove of the slider chest can synchronize if simultaneously sounded. The frequency difference between the two pipes disappears (mitnahme-effekt) giving up on a common frequency which is, in general, slightly higher than both natural frequencies of the pipes [Angster *et al.*, ''Coupling between simultaneously sounded organ pipes,'' AES preprint 3534 (F1-2), Berlin (1993)]. A related phenomenon is observed when the separation between the two natural frequencies is bigger than 1%: this gives up on an increased difference frequency. This paper shows that a numerical simulation, where the two coupled pipes are modeled with a pair of coupled nonlinear oscillators, qualitatively obtains both these results. Furthermore the existence of a parameter range fitting all the experimental data is presented.

Session 4aPA

Physical Acoustics: Propagation and Mixed Physical Topics

Roger M. Waxler, Chair
National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677

Contributed Papers

8:00

4aPA1. A simple procedure for the selection of singular values in acoustical travel-time tomography. Amin Bassrei (Inst. of Phys. and CPGG/UFBA, Campus Univ. Ondina, 40210-340 Salvador BA, Brazil)

Inverse problems are in general ill-posed and acoustical travel-time tomography is no exception. One common way to solve this inverse problem is by the generalized inversion through singular value decomposition (SVD). The crucial issue that arises here is the avoidance of small singular values which generally contribute to inconsistent solutions. The selection of singular values used in this work is based on the analysis of some estimators: rms error between the true model and the estimated one, rms error between the observed data and the calculated one, rms of the model, and the entropy of the model. In a numerical simulation with N equations and N unknowns, the SVD technique provides N singular values; N inverse matrices are then built, using $1,2,\ldots,$ up to all the N singular values. For each matrix one solution is obtained and the above estimators are calculated and analyzed. Solutions with, for instance, small rms error are accepted, and solutions with inconsistent entropies are rejected. Thus, it is possible to determine the number of singular values related to the solution with a maximum of information and a minimum of discrepancy. [Work supported by CNPq and PGS.]

8:15

4aPA2. Greens function modified spherical harmonic analysis. John MacGillivray and Victor Sparrow (Grad. Prog. in Acoust., Penn State Univ., 217 Appl. Sci. Bldg., University Park, PA 16802, creeve@sabine.acs.psu.edu)

In an effort to develop a model for impulse responses in reverberant spaces, techniques for source radiation modeling were explored. One method of decomposing a source field involves an analysis based on spherical harmonic expansions of the field. Spherical harmonic decomposition is a powerful tool, but it is a mathematical approach which is completely independent of the type of source that creates the field. As a consequence, this technique cannot, in general, be used to gain insight into the underlying physical nature of the source. Moreover, because this is a brute force technique, the analysis can be extremely inefficient. In this presentation, modifications to the spherical harmonic analysis and synthesis equations are made to overcome these problems. These are based on the (theoretical) source types associated with the individual spherical harmonic expansion coefficients, and determination of the location of the primary components of the (physical) source. These modifications can greatly improve the computational efficiency and allow the source to be described by a set of coefficients which is more directly related to actual source types present. Illustrative examples will be shown. [Work supported by the Applied Research Laboratory at Penn State University under the Educational and Foundational Research Program.]

8:30

4aPA3. Infinite impulse response-based simulation of room acoustics. Parita Amin, Jing Tsui, and Charles Thompson (Ctr. for Adv. Computation and Telecommun., ECE, Univ. of Massachusetts, Lowell, Lowell, MA 01854, Charles_Thompson2@uml.edu)

In this work, the problem of filter construction for the point-to-point impulse response in a room is examined. Our particular interest is infinite impulse response (IIR) filters. It will be shown that such filters can be constructed from image-based models. The use of subband IIR filter models is investigated. Our model provides perfect reconstruction of early arrivals. The late arrivals in the impulse response are in agreement only in the least-squared sense. Experimental results will be given for the room impulse response.

8:45

4aPA4. On the estimations of sound field by source-distributed methods. Gee-Pinn James Too (Dept. of Naval Architecture and Marine Eng., Natl. Cheng Kung Univ., Tainan, Taiwan 70101, ROC)

Source-distributed methods were developed earlier by Stepanishen *et al.* and many others to study the radiation sound field. There are several articles discussing the properties of these methods, although there are alternative names for the source-distributed methods. In the present study, the properties of these methods are summarized. The applications of source-distributed methods for radiation and scattering sound fields are demonstrated. It is interesting to show that the errors in the near-field measurement might be canceled out in the estimation in the far field. In the application for estimation of interior sound field, the source-distributed methods also give good results compared with the results of the Helmholtz integral formulation. In addition, it is pointed out that there are some restrictions in the methods. [Work supported by NSC of ROC.]

9:00

4aPA5. Application of random walk methods to wave propagation. Bair Budaev and David Bogy (Dept. of Mech. Eng., Univ. of California, Berkeley, #1740, Berkeley, CA 94720-1740)

We present an approach to problems of diffraction that has its roots in a number of well-established theories such as the geometric theory of diffraction, the methods of Feynman's path integrals and Wiener's functional integrals, and the theory of stochastic processes. We start our analysis of the Helmholtz equation by closely following the scheme of the ray method, but instead of approximating the resulting second-order auxiliary equation by a first-order transport equation, we study the original auxiliary equation and obtain its exact solution as the mathematical expectation of a specified functional in the space of Brownian trajectories with the Wiener probabilistic measure. The obtained solution appears to be a direct improvement of the ray method approximation to the exact solution of the Helmholtz equation, and it is shown to admit efficient numerical evaluation. [Work supported by NSF Grant 9730779.]

9:15–9:30 Break

9:30

4aPA6. On the square-root distance dependence of sound field near an edge of a semi-infinite plane. Mitsuhiro Ueda (Dept. of IDE, Tokyo Inst. of Technol., 2-12-1 O-okayama, Meguro-ku, Tokyo, 152-8552 Japan, ueda@ide.titech.ac.jp)

It is known that the 2-D sound field near an edge of a semi-infinite plane depends on the square-root of a distance between the observation point and the edge. The boundary value problem for the potential on the semi-infinite plane is formulated using a new principle of diffraction, that is, virtual discontinuity principle of diffraction (VDPD). According to this formulation, a new response function, that describes mutual dependence of potentials on the plane, is obtained and is localized near the edge. Since the square-root distance dependence also appears only near the edge, it becomes clear that the dependence can be explained by the analysis localized near the edge. Consequently the boundary value integral is carried out for a very short interval including the edge assuming that the observation point is very near the edge. The square-root dependence arises from the restriction that the same potential is reconstructed from the integration over the short interval. It is interesting to note that the new response function contains a scheme in it to fix the distance dependence of potential at points near the edge.

9:45

4aPA7. On the model of propagation of high-velocity subsurface waves. Edouard G. Nesvijski (Ctr. of Technol., Federal Univ. of Santa Maria, Campus Universitaio, Santa Maria, RS 97110-970, Brazil)

There are a great number of theoretical and experimental works published on the problem of analysis and classification of surface waves, but some recent experimental results cannot be explained unambiguously by existing basic theories because the velocity of registered waves is considerably higher than expected. For a better understanding of this phenomenon it is proposed to take a fresh look at the classical elaboration of the Rayleigh problem. This paper considers additional real roots of the Rayleigh equation in an attempt to explain the existence of a special type of high-velocity leaky waves propagating under the surface of half-space solid bordering with vacuum or gas. Their main energy concentrates in a local zone under the surface, and below this zone they transform into bulk waves taking away the wave energy. Some computer simulations and calculations are presented as a theoretical explanation of the conditions of high-velocity leaky subsurface waves propagation. An original representation of the Rayleigh function in 3-D gives one the opportunity to better understand the physical nature of the subsurface wave behavior in solids with different elastic characteristics. The proposed model of propagation of high velocity subsurface waves fits well with experimental data.

10:00

4aPA8. Stress measurement using critically refracted longitudinal waves. Don E. Bray (Don E. Bray, Inc., P.O. Box 10315, College Station, TX 77842)

Critically refracted longitudinal waves (LCR) offer considerable advantage in achieving nondestructive measurement of stress in engineering components and structures. Of the various wave types, the LCR is most sensitive to stress and least sensitive to texture. Also called the surface skimming longitudinal wave (SSLW), the LCR wave skims along just underneath the surface of a material at bulk wave speeds. Further, there is experimental evidence of frequency dependence where the higher frequencies concentrate their energy more toward the surface than do the lower frequency waves. This effect, which has not been fully explained theoretically, enables stress gradient evaluation below the surface. Further, there is observed scatter in the travel times found at low stress levels. At higher levels of stress, the data often become very linear. The specific nature of the LCR wave, acoustoelasticity, LCR probes and beam patterns, and methods for its excitation, are described. Probe frequencies from 1.0 to 20 MHz have been used. System travel-time repeatability is 0.01% or better and stress resolution of ± 21 MPa (3 ksi) is obtained in typical applications. Tests showing the frequency dependence and the scatter observed at low levels will be discussed. Several applications described using the LCR technique for stress measurements include pressure vessels, piping, welds, titanium turbine engine blades, aluminum plate and a turbine rotor. Future applications and research using the technique will be mentioned.

10:15

4aPA9. Analysis of fields and forces on particles in an open-sided acoustic resonator using FEM with a perfect matched layer. Yibing Zheng and Robert E. Apfel (Dept. of Mech. Eng., Yale Univ., New Haven, CT 06520-8286, yibing.zheng@yale.edu)

The standing acoustic field in an axisymmetric open-sided resonator with a curved reflector is studied using the finite element method (FEM) with an anisotropic perfectly matched layer (PML) boundary condition. This kind of open-sided resonator is widely used in acoustic and acoustoelectric levitators, where the acoustic field is not considered as one-dimensional due to the shape of the reflector. The method has been successfully employed to calculate the acoustic field profiles, the resonance frequencies, and the quality factors of the resonant modes. The numerical results of resonators with a concave or convex reflector are compared with those of the analytical asymptotic method. In the resonator of a real acoustic levitator, this method can calculate the acoustic fields driven by the surface vibration of the transducer, with and without particles levitated. The acoustic radiation forces on levitated particles are analyzed in such acoustic standing fields. The authors are trying to extend this study to characterize the levitation of a three-dimensional drop array in an acoustoelectric levitator. [Work supported by NASA (NAG3-2147) and by NSF (CTS-9870015).]

10:30

4aPA10. The effect of time-dependent void fractions in the oscillations of wet foam drops. Iskander Akhatov (Inst. of Mech., Ufa Branch of the Russian Acad. of Sci., 6 Karl Marx St., Ufa 450000, Russia, iskander@anrb.ru), J. Gregory McDaniel, and R. Glynn Holt (Boston Univ., 110 Cummington St., Boston, MA 02215)

We develop a model for the nonlinear oscillations of spherical drops composed of aqueous foam. Beginning with a simple mixture law, and utilizing a mass-conserving bubble-in-cell scheme and a uniform pressure approximation, we obtain a Rayleigh–Plesset-type equation for the dynamics of bubbles in a foam mixture. The dispersion relation for sound waves in a bubbly liquid is then coupled with a normal-modes expansion to derive expressions for the frequencies of modal oscillations. These modal (breathing plus higher-order shape modes) frequencies are elicited as a function of the void fraction of the foam. A Mathieu-type equation is obtained for the dynamics of the higher-order shape modes and their parametric coupling to the breathing mode. Numerical results are presented. [Work supported by NASA.]

10:45–11:00 Break

11:00

4aPA11. A study of knock and adiabatic process in multi-fuel engines. Salvador Cerdá, José Romero, Joaquín Navasquillo (Lab of Acoust., Dept. of Appl. Phys., Univ. of Valencia, Spain), and Grover Zurita (Lulea Univ.)

For both engine designers and control purposes, a better understanding of the knocking that occurs during combustion in engines can lead to optimal decisions which will make engines operate more efficiently. The rapid rise in diesel engines is recognized as an audible impulse noise, which is known as ''knock.'' The knock is caused by the spontaneous combustion of a significant volume of fuel/air mixture. The knock produces shock waves within the combustion chamber and results in a great number of combustion chamber resonance frequencies. The engine used in the experiments was a six-cylinder modified diesel engine running in ethanol conditions. The main purpose of this research is to determine the

knock phenomena and the adiabatic process. To perform the data analysis a time-frequency spectral approach was used. The results revealed both radial and circumferential resonances. The most important adiabatic coefficients were carried out and used to calculate the heat release rate.

11:15

4aPA12. Sound attenuation in the critical region of the ionic binary mixture ethylammonium nitrate/*n*-octanol. Sirojiddin Z. Mirzaev (Heat Phys. Dept., Uzbek Acad. of Sci., Katartal 28, Tashkent 700135, Uzbekistan) and Udo Kaatze (Drittes Physikalisches Institut, Georg-August-Universitaet, 37073 Goettingen, Germany)

Ultrasonic attenuation spectra (100 kHz$<f<$4 MHz) of the ionic mixture ethylammonium nitrate/*n*-octanol of critical composition and also of the ethylammonium nitrate itself (100 kHz $<f<$ 400 MHz) are discussed at different temperatures. The measured spectra are evaluated in terms of the Bhattacharjee–Ferrell dynamic scaling theory. Using literature data for the amplitude of the fluctuation correlation length, the background and critical part of the heat capacity, the sonic attenuation spectrum as predicted by the Bhattacharjee-Ferrell model has been calculated for the critical mixture at the critical temperature. Again following this theoretical model, the contribution due to concentration fluctuations has been subtracted from the measured spectra at the temperatures of measurement and also subtracted has been the high-frequency asymptotic background contribution. The measured data are consistent with the Bhattacharjee–Ferrell theory of critical concentration fluctuations but show an additional frequency dependence. [Work supported by the Volkswagen-Stiftung, Hannover, Germany.]

11:30

4aPA13. Results of attenuation calculations for a three-component gas mixture. Yefim Dain and Richard M. Lueptow (Dept. of Mech. Eng., Northwestern Univ., 2145 Sheridan Rd., Evanston, IL 60208, r-lueptow@northwestern.edu)

Acoustic attenuation in a mixture of gases results from the combined effects of molecular relaxation and the classical mechanisms of viscosity and heat conduction. As a result, the attenuation depends on the composition of the gas, frequency, temperature, and pressure. A model of the relaxational attenuation that permits the prediction of acoustic attenuation is used to predict the effect of composition, frequency, temperature, and pressure on the acoustic attenuation in a three-component gas mixture of nitrogen, methane, and water vapor. The attenuation spectrum is dependent upon the composition through the appearance of peaks in the spectrum related to the relaxation frequency of the particular components and their relaxing complexes. The relaxation peak related to methane dominates the spectrum except at low methane concentrations, where the nitrogen peak, which is dependent upon the water vapor concentration, is evident. Temperature and pressure significantly alter the relaxation frequency and the degree of attenuation, but water vapor plays little role in the attenuation at moderate and high methane concentrations. [This work was funded by the Department of Energy under subcontract from Commercial Electronics, Inc. (Broken Arrow, OK) and by the Ford Motor Company.]

11:45

4aPA14. A study of simulation methods of limited diffraction beams. Hu Peng, Zuosheng Zhang (Univ. of Sci. Tech. China, Hefei, PROC 230026), and Jian-yu Lu (Univ. of Toledo, Toledo, OH 43606)

Limited diffraction beams such as Bessel beams and *X* waves have been studied extensively because of their potential applications in many areas such as medical imaging and optical communications. Computer simulations have been powerful tools for understanding the fundamentals of these beams. In this study, limited diffraction beams, specifically, Bessel beams and *X* waves, were simulated. The simulations were based on the frequency-domain method (Rayleigh–Sommerfeld diffraction formula), time-domain method, and finite-element or finite-difference method. For Bessel beams, a circular transducer divided into multiple rings was assumed. For pressure distributions on axial axis, the simulations were simplified due to the symmetry. In the case of *X* waves, the differential property of acoustic waves was used. The results of the simulation methods were compared with each other. Advantages and disadvantages of these methods will be reported. In all cases, the results are very close to those predicted by theory. [This work was supported in part by Grant HL 60301 from the National Institute of Health of the U.S.A.]

THURSDAY MORNING, 7 DECEMBER 2000 CALIFORNIA SALON 2, 8:30 A.M. TO 12:00 NOON

Session 4aPP

Psychological and Physiological Acoustics: Pitch, Loudness and Localization

Kourosh Saberi, Chair
Department of Cognitive Sciences, University of California, Irvine, California 92697

Contributed Papers

8:30

4aPP1. Franssen à la phones. William M. Hartmann, Scott R. Lawton (Phys. and Astron., Michigan State Univ., East Lansing, MI 48824), and Brad Rakerd (Michigan State Univ.)

According to current explanation, the Franssen effect reflects a competition between localization cues from the onset of a tone and cues from the steady-state portion. For a sine tone in a room, the steady-state interaural time and level cues (ITD and ILD) are implausible because of room reflections. Therefore, onset cues dominate, and a listener localizes the tone at its onset position regardless of the true origin of the steady state. To test this explanation we used headphones to simulate the simplest kind of Franssen stimulus: The ITD was fixed throughout the tone. At onset, the ILD was zero. After onset, the ILD grew so as to contradict the ITD in the steady state. Equal numbers of trials were done with abrupt and slow onsets. The listener's task was to localize the steady-state part of the tone. A Franssen analog was identified by comparing responses to abrupt-onset and slow-onset tones. Results were compared with the standard Franssen effect in a large room. Clear evidence of a Franssen effect, with its dependence on plausibility, appeared in the headphones experiments, but the data were less dramatic than data obtained in the room. [Work supported by the NIDCD.]

8:45

4aPP2. Specific learning phenomena in pitch discrimination tasks. Laurent Demany, Catherine Semal, Cecilia Maubaret, and Odon Noblia (Lab. de Neurophysiol., CNRS and Univ. Victor Segalen, 146 rue Leo Saignat, F-33076 Bordeaux, France, Laurent.Demany@psyac.u-bordeaux2.fr)

In experiment 1, listeners were trained to discriminate, during ten sessions (11 000 trials), a monaural pure tone of 3000 Hz from ipsilateral pure tones with slightly different frequencies. The resulting perceptual learning (improvement in thresholds) appeared to be frequency-specific since, in subsequent sessions, a new learning was observed when the tone frequencies were shifted to a range near 1200 Hz, or 6500 Hz. By contrast, when the initial tones were subsequently presented to the contralateral ear, thresholds were immediately optimal; therefore, the initial learning was not ear-specific. In experiment 2, performed on different listeners, training in pitch discrimination was initially provided using complex tones that consisted of harmonics 3–7 of a missing fundamental (near 100 Hz for some listeners, 500 Hz for others). After 11 000 training trials, these complex tone stimuli were replaced by pure tones with frequencies that compared to the complex tone frequencies in three different ways: (condition A) their missing fundamentals, (condition B) their median spectral components, (condition C) completely novel. New learning was observed in each condition. Although the pure tones used in condition A were matched in pitch to the previously presented complexes, there was no evidence for a transfer of learning in this condition.

9:00

4aPP3. The upper boundary of temporal pitch coding inferred from electric stimulation of the auditory nerve. Fan-Gang Zeng (Univ. of California, 364 Med Surge II, Irvine, CA 92697)

Pitch can be represented by a vibration pattern in the cochlear partition (the place code) or a spike firing pattern in the auditory nerve (the temporal code). Generally, pitch of high tones has been assumed to use the place code and pitch of low tones to use the temporal code. Because both place and temporal codes change simultaneously as a function of frequency in a normal auditory system, the upper boundary of temporal encoding of pitch is not clear. Here, we use cochlear implants to infer the upper boundary of temporal pitch coding. The cochlear implant electrically stimulates the auditory nerve, providing an opportunity to assess the temporal pitch coding independent of the place code. Cochlear implant listeners performed pitch estimate and discrimination tasks as a function of frequency. The tasks were performed when either basal or apical electrodes were stimulated independently. The results showed that pitch is monotonically increased for frequencies lower than 300 Hz and does not change for frequencies above that. Pitch discrimination results also showed a break point at about 300 Hz. Together, the present results suggest that 300 Hz is the upper boundary of temporal pitch coding.

9:15

4aPP4. Testing a new theory of Fechnerian scaling: The case of auditory intensity discrimination. Karin Zimmer and Hans Colonius (Inst. fur Kognitionsforschung, FB5/A6, Universitat Oldenburg, 26111 Oldenburg, Germany)

Recently, Dzhafarov and Colonius [Psych. Bull. Rev. **6**, 239–268 (1999)] proposed a theory that refines Fechner's ideas of deriving the perceptual distance between stimuli from discriminability measures in the unidimensional case, and generalizes it to multidimensional stimulus domains. In this theory, the existence of a Fechnerian metric is based on testable assumptions about the shape of the psychometric functions. Thus, ''same–different'' psychometric functions: (1) have to attain a single minimum at some point; (2) must be locally symmetric around their minima; and (3) must show the same degree of cuspidness, or flatness, along the stimulus continuum. In the present investigation, these assumptions were evaluated experimentally employing acoustical stimuli differing along one stimulus dimension, intensity. In a ''same–different'' paradigm, using the method of constant stimuli, seven 1000-Hz tones deviating from 0 to 2.1 dB with respect to a standard were presented diotically to five observers. For every participant, psychometric functions were assessed at five hearing levels ranging from 40- to 80-dB HL. The outcome of the ''same–different'' experiments generally supports the assumptions of the theory. Therefore, data such as these are, in principle, suitable for deriving a global loudness scale from local discriminability measures.

9:30

4aPP5. Intensity-difference limens as a function of level for Gaussian-shaped packets of broadband noise. Lance Nizami, Walt Jesteadt, and Jason F. Reimer (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131, nizami@boystown.org)

The nonmonotonic level dependence of the intensity-difference limens (DLs) for a broadband stimulus, the click [Raab and Taub, J. Acoust. Soc. Am. **46**, 965–968 (1969)], can also be observed for a narrow-band stimulus, a Gaussian-enveloped 2-kHz tone [Nizami *et al.*, J. Acoust. Soc. Am. **107**, 2820 (2000)]. Although the DLs for a 32-ms tone-pip (measured at plus-or-minus four standard deviations) follow the near-miss to Webers Law, the DLs for a 4-ms tone-pip are higher, and peak mid-level, at about 40 dB SL. The DLs are larger still, and the peak more profound, for an 8-ms tone-pip. Energy-spectrum changes as duration changes, however, obscuring the source of the nonmonotonicity. Energy-spectrum was therefore held constant by using a broadband-noise carrier, while preserving the Gaussian envelope. DLs were obtained from two male and four female subjects (ages 18–40) having normal hearing thresholds. Noise-pip levels ranged from 30 to 90 dB SPL at durations of 2, 4, 8, 16, and 32 ms. Although the across-subject mean DLs for 32- and 16-ms noise-pips follow Weber's law, the means for 8 ms are larger, and rise at mid-level. The DLs grow as duration decreases from 8 to 4 to 2 ms, unlike the case for 2 kHz.

9:45

4aPP6. Temporal weights for loudness comparisons. Wolfgang Ellermeier and Sabine Schrödl (Institut für Psychologie, Universität Regensburg, 93040 Regensburg, Germany)

Temporal integration of the loudness of level-fluctuating sounds was investigated using the trial-by-trial analysis developed by Berg [J. Acoust. Soc. Am. **86**, 1743–1746 (1989)]. In a two-interval, forced-choice procedure, two noise samples of 1-s duration were presented on each trial. Their levels were randomly varied every 100 ms by repeatedly, and independently, drawing from normal distributions having expected values of 60 or 61 dB SPL, respectively. The subject's task was to decide which of the two noise samples sounded louder. Based on a few thousand judgments, conditional psychometric (COSS) functions were estimated in order to determine to what extent each temporal segment contributed to the judgment. Data collected on six subjects showed considerable individual differences in the weighting functions, but—in contrast to averaging techniques widely used in sound evaluation—exhibited a statistically significant trend to accentuate the onset of the noise. These results were compared to the outcome of a second experiment, in which the total duration of the noise samples was shortened to 200 ms to address a different stage of echoic memory. Conclusions regarding the sensory and cognitive mechanisms involved in short-term loudness integration will be presented.

10:00–10:15 Break

10:15

4aPP7. Effect of frequency region and noise bandwidth on the AMFR. Shaum P. Bhagat and Craig A. Champlin (Univ. of Texas, Austin, Austin, TX 78712)

This study investigated the effect of varying stimulus frequency region and noise bandwidth on the amplitude modulation following response (AMFR). The stimulus for the AMFR was an amplitude-modulated burst of bandlimited white noise. Two carrier frequency regions (1 and 4 kHz) and two noise bandwidths (100 Hz and 1 kHz) were employed. Three women and three men with normal hearing participated in the study. Low carrier-frequency regions produced significantly larger AMFRs than high

carrier-frequency regions. Wide-noise bandwidths generated significantly larger AMFRs than narrow-noise bandwidths. The noise bandwidth effect was dependent on the rate of amplitude modulation. Modulation transfer functions were derived by plotting the AMFR amplitude as a function of modulation rate. At the greatest modulation depths, the transfer function resembled a bandpass filter. For lower modulation depths, the transfer function resembled a low-pass filter. The results of the study are in general agreement with previous research that investigated the AMFR with tone and wideband noise carrier signals.

10:30

4aPP8. Distance localization using nonindividualized head-related transfer functions. Pavel Zahorik (Dept. of Psych. and Waisman Ctr., Univ. of Wisconsin—Madison, Madison, WI 53705)

Past research has suggested that virtual auditory displays are most accurate when implemented using individualized head-related transfer functions (HRTFs). Directional localization accuracy, for example, is often substantially degraded when nonindividualized HRTFs are used, due to the resulting distortion of spectral cues to sound source direction. Although equally important for accurate spatial representation, little is known about distance localization accuracy in virtual auditory displays, and specifically, whether individualized HRTFs are necessary for accurate distance localization. Here, distance localization accuracy was assessed for six listeners, using both individualized and nonindividualized HRTFs in separate blocks of trials. The virtual sound sources were produced from special HRTF measurements (source distance ranging from 0.3 to 13.8 m) that also contained the acoustic transfer characteristics of a semireverberant room environment. This provided listeners with additional acoustic cues to source distance, including the ratio of direct-to-reverberant sound energy. Since no significant differences were observed in listener's judgments of apparent sound source distance in the two stimulus conditions, it is concluded that distance localization accuracy, and degree of sound source externalization, do not depend on the use of individualized HRTFs when natural reverberation is available to the listener. [Work supported by NIH-NIDCD, NASA, and NIH-NEI.]

10:45

4aPP9. Reverberant speech intelligibility for monaural and binaural listening. Brad Libbey and Peter Rogers (Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405, gt1556a@prism.gatech.edu)

Humans listening to speech in a small room are frequently unaware of reverberation. It is unknown if neurological processes remove these echoes or if they are simply disregarded when speech is phonetically processed. In other words, is there a neurological mechanism that is capable of removing echoes to create a clean speech neurological signal before phonetic processing? Or is the brain capable of processing reverberant phonemes? These questions have been explored in the process of investigating the hypothesis; ''Binaural listening does not improve speech intelligibility.'' Previous experiments have shown that binaural listening does not improve intelligibility for highly reverberant small rooms, T60 $=1.5$ s. The current experiment investigates the intelligibility of words in three reverberant rooms, T60$=1.1$, 0.7, and 0.4 s. Subjects listen with one or two ears to lists of words taken from a standardized testing procedure, ASA 85. In addition the rooms are simulated using the method of images and the reverberant words are presented in a diotic and binaural format over headphones. Intelligibility results from these rooms with monaural and binaural listening are presented.

11:00

4aPP10. A model to simulate human sound localization and detection in presence of a distracter. Jonas Braasch (Institut für Kommunikationsakustik, Ruhr-Universität, Bochum, Germany)

To simulate human auditory localization performance of a target in the presence of a distracter, existing models focus the analysis of the targets on frequency bands or time gaps where the distracter is not present. However, human localization also works quite well if the frequency ranges of target and distracter are identical and the distracter is played during the complete duration of the target—as long as the onsets of target and distracter are separated in time. This is taken into account in the proposed model. For each frequency band, the interaural time differences of the target are calculated from the difference of the interaural cross correlation (ICC) of the total sound (target+distracter) and the ICC of the distracter alone ($\Phi_T = 3D\Phi_{TD} - \Phi_D$, assuming that target and distracter are uncorrelated). The latter is calculated from the part of the distracter that precedes the target. A similar procedure is used to measure the interaural level differences of the target. Auditory events are simulated assuming an ideal observer and internal noise. The model is able to mimic a number of psychoacoustic effects, including spatial shifts of the auditory events for low SNR situations. [Work supported by DFG.]

11:15

4aPP11. Individualized HRTFs using computer vision and computational acoustics. Ramani Duraiswami, Larry Davis, Shihab A. Shamma, Howard C. Elman (Inst. for Adv. Computer Studies, Univ. of Maryland, College Park, MD 20742), Richard O. Duda, V. Ralph Algazi (Univ. of California, Davis, Davis, CA 95616), Qing-Ho Liu (Duke Univ., Durham, NC 27708), and S. T. Raveendra (Automated Anal. Corp., Ann Arbor, MI 48104)

For a given sound source location and frequency content, the sound received at the two ears of a person is first scattered by his/her torso, head, and pinnae, leading to interaural differences in the intensity and spectral features of the sound received. These features are used as cues to perceive the spatial location of the source, and are encoded in an extremely individual ''head-related transfer function'' (HRTF) that depends on the person's external anatomy (structure of the torso, head, and pinnae). This individuality has made it difficult to use the HRTF to render perceptually convincing spatially distributed audio for applications such as entertainment or virtual reality. We have commenced a long-term project that will use numerical methods to compute individualized HRTFs from accurate 3D surface models of a person's torso, head, and pinnae (obtained using computer vision techniques). These surface models will be used with boundary-element methods to solve the wave equation and compute the HRTFs. An overview of our long-term project, as well as preliminary results that compare the numerical techniques to be used with analytical and experimental results for scattering from simple shapes, will be presented. [Work supported by NSF.]

11:30

4aPP12. AM depth versus degree of amplitude fluctuation: Implementation error, adjustment, and implications. Pantelis Vassilakis (Systematic Musicology, Music Percept. and Acoust. Lab., UCLA, Box 951657, Los Angeles, CA 90095-1657)

AM depth values are often used in signal processing as a measure of amplitude fluctuation. In such cases, AM implementation produces an error reflecting the nonlinear relationship between the presumed and applied AM depths. A distinction has been made between two concepts of AM depth: (a) ''spectral AM-depth,'' referring to the relationship between AM depth and degree of spectral energy spread and (b) ''envelope AM-depth,'' referring to the relationship between AM depth and degree of amplitude fluctuation. The usual AM implementation equation represents consistently only the ''spectral AM-depth.'' It is shown that, in order to apply an intended ''envelope AM-depth,'' an adjusted AM depth coefficient must be inserted in this equation. The adjustment affects studies where changes in the degree of amplitude fluctuation have been correlated to changes in some physical, physiological or perceptual measurement. In general, all functions derived from the results of studies based on the miscalculation of amplitude fluctuation have been distorted by a factor proportional to the identified error. If and how this distortion influenced the interpretation of the provided functions will vary according to the importance and meaning attributed to each specific functions shape. Specific examples and general implications are discussed.

11:45

4aPP13. Audiometric tests with various earphones, signals of short duration, and signals of high frequencies. Morten Lydolf, Brian L. Karlsen, Arturo O. Santilln, and Henrik Moller (Dept. of Acoust., Aalborg Univ., Fredrik Bajers Vej 7B-4, DK-9220 Aalborg O, Denmark, acoustics@acoustics.dk)

New audiometric standards on tests with signals of short duration and tests with signals of high frequencies are in the course of preparation by the international working group, ISO/TC 43/WG 1. In this paper, first a selection of five different audiometric earphones is compared by using threshold measurements in the frequency range 125 Hz–8 kHz. These data are meant for validation of the reference threshold for regular pure-tone audiometry given in ISO 389-1. Signals of short duration consist of either rectangular pulses with a duration of 0.1 ms or pure-tone samples including only seven full periods at the octave band frequencies 0.5–8 kHz. Both single events and bursts with 20 signals presented during 1 s were used as stimuli. Finally, audiometric tests were made with high-frequency pure-tone stimuli of 1-s duration at: 8, 9, 10, 11.2, 12.5, 14, and 16 kHz. For all experiments the chosen psychophysical method was the ascending method, implemented nearly as described in ISO 8253-1. Twenty-five young, otologically normal test subjects were used in order to contribute to the standardization work. Representative reference thresholds are needed with respect to signals of short duration for the coming ISO 389-6 and for high frequencies in ISO 389-5.

THURSDAY MORNING, 7 DECEMBER 2000 CALIFORNIA SALON 5, 8:25 A.M. TO 12:00 NOON

Session 4aSA

Structural Acoustics and Vibration: Acoustic Microsensors

Vasundra V. Varadan, Chair
212 Earth Engineering Sciences Building, Pennsylvania State University, University Park, Pennsylvania 16802

Chair's Introduction—8:25

Invited Papers

8:30

4aSA1. Acoustic atomic force microscopy applications in nanotechnology. Bernhard R. Tittmann, Weilou Gao, and Jikai Du (Dept. of Eng. Sci. and Mech., Penn State Univ., University Park, PA 16802)

In this work, acoustic atomic force microscopy is presented for nanometer material characterization. A transducer with a frequency range from 0.1 kHz to 1 MHz is typically fixed under a sample. The resonant frequency for the composite system is determined experimentally. The amplitude and phase images of the vibrating composite are obtained simultaneously as well as surface topography. Plasm spray drying process is studied by use of the acoustic AFM. Enhanced contrast in amplitude images shows unique capability to determine the plasm distribution in Alumina matrix. Further study is conducted on the contrast mechanism with a sample comprised of three different materials: Si3N4, Cu, and steel. Theoretical analysis based on finite-element methods (FEM) and Johnson, Kendall, Roberts (JKR) model is conducted to simulate the system and determine the resonant frequency, amplitude, and phase. The results of the numerical simulations compare favorably with the experiment results. Furthermore, an amplitude-based method is proposed for estimating the material modulus through a semitheoretical, seminumerical process. The method shows higher sensitivity than the resonant frequency-based method. Also, microelectronic devices are studied for the analysis of diffusion.

9:00

4aSA2. Silicon micromachined ultrasonic transducers for bulk, Lamb, and Rayleigh waves. B. T. Khuri-Yakub (E. L. Ginzton Lab., Rm. 11, Stanford Univ., Stanford, CA 94305-4085, khuri-yakub@stanford.edu)

Capacitive micromachined ultrasonic transducers (cMUT) can match and even outperform piezoelectric transducers in terms of efficiency and bandwidth. With the advent of silicon micromachining, it is now possible to make capacitors with very thin gaps that sustain electric fields of the order of 109 V/m or more. At these levels of electric field, the transformer coupling between the electrical and mechanical parts of the capacitor transducer, and thus its performance, are comparable to that of piezoelectric transducers. We will review the design and performance of cMUTs for airborne, immersion, Lamb and Rayleigh waves' ultrasound applications. We will show examples of immersion transducers (single element, 1-D and 2-D arrays of elements), air transducers and Lamb wave transducers. We will also present transducer systems (single element, 1-D and 2-D arrays of elements) with dynamic ranges of over 150 dB/Volt/vHz, which are in excellent agreement with the theoretical predictions.

9:30

4aSA3. Wireless surface acoustic wave and MEMS-based microsensors. Vijay K. Varadan (Penn State Univ., 212 Earth/Eng. Sci. Bldg., University Park, PA 16802)

The integration of surface acoustic waves (SAW), microelectromechanical systems (MEMS), and required microelectronics and conformal antennas to realize programmable microsensors suitable for many engineering and biomedical applications will be reviewed in this talk. This unique combination of technologies results in novel conformal sensors that can be remotely sensed by a wireless communication system with the advantage of no power requirements at the sensor site (passive sensor). The required features in many

of these applications are high precision, wide dynamic range, and wide frequency range. The MEMS-SAW-based devices presented possess typical advantages of MEMS sensors including the additional benefits of robustness, excellent sensitivity, surface conformability, and durability. After a brief overview of SAW sensors and MEMS, the paper is focused on the design and fabrication of MEMS devices for engineering and medical applications, for example: accelerometer and gyroscopes for automobile, inertial navigation; sensors for pressure, torque, temperature, and strain; health monitoring of structures; ''smart tongue and electronic nose;'' micropump for drug delivery; humidity, temperature, and ethlyne oxide sensor for hospital sterilization system; microsensors for monitoring bone healing; and pollution and chemical sensing.

10:00–10:15 Break

10:15

4aSA4. Finite-element method for determination of electromechanical coupling coefficient for piezoelectric and capacitive micromachined ultrasonic transducers. John D. Fraser (ATL Ultrasound, a Philips Medical Systems Co., 22100 Bothell Everett Hwy., P.O. Box 3003, Bothell, WA 98041-3003, john.fraser@philips.com) and Paul Reynolds (Weidlinger Assoc., Inc., Los Altos, CA 94022-1049)

Research has been reported on ultrasonic transduction using capacitive micromachined ultrasonic transducers (cMUTs). These are thin membranes of diameters of order 100 μm suspended over a silicon substrate with a gap of order 1 μm. The devices become transducers when a bias voltage is applied, causing electrostatic forces to draw the membranes closer to the substrate. Operation is by applying an electrical signal to excite a vibration in the membrane, or an ultrasonic wave to excite an ac voltage. The cMUT literature includes equivalent circuits and studies on losses and spurious waves, and discussion of output, sensitivity, and bandwidth. Transducer engineers would like to see analogies to the properties of piezoelectric transducers, in order to compare the cMUTs to traditional designs. An extension of the definition of the piezoelectric coupling coefficient to an effective coupling coefficient for cMUT transducers is presented. A finite-element method has been developed using the PZFLEX software package to determine its value, and used for evaluating cMUT designs. Accepted values are obtained for a piezoelectric material, and it is shown that an ideal cMUT has a coupling coefficient approaching 1 at the membrane collapse voltage, while practical devices are often limited to lower values.

10:45

4aSA5. Computational modeling and analysis of SAW devices. Guanshui Xu (Dept. of Mech. Eng., Univ. of California at Riverside, Riverside, CA 92521)

Direct full-scale computational simulation of electronic acoustic wave devices is among the most challenging problems for designing high performance devices. The problem also becomes urgent as the conventional methods fail to characterize the behavior of miniaturized devices that operate at high frequency for demanding applications. The effective computational methodology has not yet been developed for the full-scale simulation of multi-physical phenomena in many micro-electromechanical devices of complex geometric structure on the nano- and micro-scale. Here a direct finite element model for the full-scale analysis of the frequency response of surface acoustic wave (SAW) devices is presented. The ability of the method to quantitatively evaluate the influence of the bulk waves at the high-frequency end of the filter passband and the influence of the number of electrodes on insertion loss is noteworthy. Some practical computational challenges of finite element modeling of SAW devices are also discussed parenthetically.

Contributed Papers

11:15

4aSA6. Development of a SAW microgas sensor for monitoring SO_2 gas. Yongrae Roh, Youngjin Lee, and Hakbong Kim (Dept. of Sensor Eng., Kyungpook Natl. Univ., Sankyug-Dong 1370, Taegu 702-701, Korea, yryong@eeg.kyungpook.ac.kr)

We have developed a new type SO_2 gas sensor applying a particular inorganic thin film on SAW devices. The sensor consists of twin SAW oscillators of the center frequency of 54 MHz fabricated on $LiTaO_3$ piezoelectric single crystals. One delay line of the sensor is coated with a CdS thin film that selectively adsorbs and desorbs SO_2, while the other is uncoated for use as a stable reference. Deposition of CdS thin film has been carried out by the spray pyrolysis method using an ultrasonic nozzle. Mass loading induced by the SO_2 gas adsorbed onto the CdS film results in corresponding frequency shifts directly proportional to the gas concentration. The relative change in the frequency of the two oscillators is monitored with a digital signal processing circuit. SO_2 sensor properties investigated include selectivity, sensitivity, response time, and repeatability. The sensor shows promising performance as a microsensing tool and is capable of measuring concentrations in air less than 200 parts per billion of SO_2.

11:30

4aSA7. Wireless surface acoustic wave ice sensor. K. A. Jose, S. Gangadharan, V. V. Varadan, and V. K. Varadan (212 Earth-Eng. Sci. Bldg., Penn State Univ., University Park, PA 16801)

Ice formation on an aircraft can cause serious safety problems, and hence, there is a real need for ice detection and monitoring on aircrafts. It is well known that acoustic wave devices offer a simple and inexpensive technique for a number of sensing applications. In this paper, we present a wireless ice detection and monitoring system using a surface acoustic wave (SAW) device utilizing shear horizontal (SH) waves as the sensing mechanism. The SAW device is designed and fabricated on a quartz substrate. Silicon dioxide is then deposited to serve as a guiding layer for SH waves. The acoustic wave concentrated in the guiding layer is highly sensitive to surface perturbations. The experimental setup used to measure the effect of ice formation is by an oscillator setup. The SAW device is designed to oscillate at a particular frequency and any fractional change in wave velocity due to surface perturbations leads to a change in oscillation frequency. It is observed that phase transformation from water to ice causes a shift in oscillator frequency of 2 MHz. The ice sensor can be operated wirelessly by connecting the SAW sensor to a stand-alone oscillator and the oscillator can communicate with a remote monitoring system using a microstrip antenna.

11:45

4aSA8. Solid-state MOSFET-based micro hydrophone array. Bei Zhu, J. Ziang, Vijay K. Varadan, and Vasundara V. Varadan (Penn State Univ., University Park, PA 16802)

The fabrication of low-cost, high-sensitivity, and miniaturized hydrophone array is highly desirable for application in sonar technology. However, as hydrophone size decreases, it becomes necessary to provide an amplifier or buffer in close proximity to avoid sensitivity loss due to interconnect capacitance. This suggests the concept of hydrpophone integration with microelectronics in an integrated circuit environment. A proptotype of an integrated transducer was thus realized in which a P(VDF-TrFE) film was bonded to the extended gate of a metal–oxide–semiconductor field-effect transistor (MOSFET) which was fabricated on a silicon wafer using standard NMOS process technology. This technique is compatible with IC fabrication processes and enables the integration of active elements with the transducer element on the same chip. The integrated structure eliminates long leads between the transducer and the amplifier and makes the net transducer–amplifier detectivity superior to the ceramic transducer–amplifier system. The devices are then encapsulated with Rho-C rubber which results in good acoustic impedance to that of water. The acoustic sensing performance was carried out in a pulse tube over a frequency range from 4 to 10 kHz. The experimental results along with the theoretical prediction by Mason equivalent circuit will be presented.

THURSDAY MORNING, 7 DECEMBER 2000 NEWPORT ROOM NORTH, 8:30 TO 11:00 A.M.

Session 4aSCa

Speech Communication: Auditory Function and Segmental Perception (Poster Session)

Ann R. Bradlow, Chair

Department of Linguistics, Northwestern University, 2016 Sheridan Road, Evanston, Illinois 60208

Contributed Papers

All posters will be on display from 8:30 a.m. to 11:00 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:30 a.m. to 9:45 a.m. and contributors of even-numbered papers will be at their posters from 9:45 a.m. to 11:00 a.m.

4aSCa1. Auditory temporal resolution and speech performance in cochlear implant users. Qian-Jie Fu (House Ear Inst., 2100 W. Third St., Los Angeles, CA 90057)

It has been widely assumed that speech recognition depends, among other factors, on the auditory temporal resolution function. The present study was conducted to investigate the relationship between auditory temporal resolution and speech performance in nine Nucleus-22 cochlear implant (CI) listeners fitting with the SPEAK strategy. A modulation detection threshold task was used to evaluate auditory temporal resolution. Modulation detection thresholds were measured as a function of current levels. Results showed some marked difference in the temporal modulation function across CI users. Some patients showed increased modulation thresholds across the entire dynamic range while others showed thresholds asymptoted as low as 30% of the dynamic range. There is a strong correlation between consonant recognition scores and the mean modulation detection thresholds (across the entire dynamic range) among CI users. The correlation between the mean modulation detection thresholds and vowel recognition scores was relatively weak but was still significant. The results suggest that phoneme recognition, especially consonant recognition, depends highly on the patients' capability to detect temporal modulation. One implication from the present results is that the high variability of speech performance is at least partly due to the differences in the temporal resolution of CI users. [Work supported by NIDCD.]

4aSCa2. Intelligibility of words by reduced-channel and cochlear implant listeners. Thomas Carrell, Kathy Shapley (Commun. Disord., Univ. of Nebraska, Lincoln, NE 68583), and Dawn Koch (Northwestern Univ., Evanston, IL 60208)

One of the differences in the way a cochlear implant, versus a normally functioning cochlea, delivers sound to auditory centers in the brain is that the implant delivers far fewer channels of frequency information. This experiment investigated how speech intelligibility in cochlear implant (CI) listeners may be related to reduced frequency resolution as opposed to other factors. A closed-response set Modified Rhyme Test was administered to CI listeners [House *et al.*, J. Acoust. Soc. Am. **37** (1965)] using naturally produced words spoken in isolation. The stimuli presented to the listeners without implants were a reduced-channel (RC) version of exactly the same stimuli, which were filtered and processed in software in a manner similar to the hardware method described by Shannon [Science **270** (1995)]. The effect of the signal processing was to divide each word into four frequency channels for the first group of RC listeners and into eight channels for the second group. Results indicated final phoneme identification appeared to be more difficult for CI listeners than for four- or eight-band RC listeners. Results were more variable for phonemes in the initial position. CI listeners were better than both four- and eight-band listeners on affricates and better than eight-band listeners for weak fricatives.

4aSCa3. Speech perception, production, and intelligibility improvements in vowel-pair contrasts among adults who receive cochlear implants. Jennell Vick, Harlan Lane, Joseph Perkell, Melanie Matthies (Speech Commun. Group, Res. Lab. of Electron., Rm. 36-511, M.I.T., 50 Vassar St., Cambridge, MA 02139), John Gould, and Majid Zandipour (Res. Lab. of Electron., Cambridge, MA 02139)

This study investigates relations among speech perception, speech production, and intelligibility in postlingually deaf adults who receive cochlear implants (CI). Measures were made for seven vowel pairs that neighbor in acoustic space from eight postlingually deafened adults, pre- and postimplant. Improvements in a speaker's production, perception, and intelligibility of a given vowel contrast tended to occur together. Subjects who produced vowel pairs with reduced contrast in the preimplant condition (measured by separation in the acoustic vowel space) and who showed improvement in their perception of these contrasts postimplant (measured with a phoneme identification test) were found to have improved production contrasts post-CI. These enhanced production contrasts

were associated with enhanced intelligibility, as measured from responses of a group of normal-hearing listeners. The results support the hypothesis that the implant user's improving speech perception contributes, at least in part, to that speaker's improving speech production. [Work supported by the NIDCD, NIH.]

4aSCa4. Channel interaction and speech processing strategies for cochlear implants. Ginger S. Stickney, Philip Loizou, Lakshmi Narayan Mishra, Peter F. Assmann (Univ. of Texas at Dallas, Box 830688, Richardson, TX 75083-0688), Robert V. Shannon (House Ear Inst., Los Angeles, CA 90057-1902), and Jane M. Opie (InSonus Medical, Inc., Newark, CA 94560)

A limitation of multichannel cochlear implants is that electrical channel interactions can arise when multiple electrodes are stimulated simultaneously. Channel interaction can be reduced by decreasing current levels delivered to each electrode, through improved electrode positioning and design, or by using speech processing strategies that stimulate electrodes sequentially. The goal of the present study was to examine the relationship between channel interaction and speech recognition for several speech strategies varying in the number of electrodes stimulated simultaneously. Based on previous research, susceptibility to channel interaction is expected to vary as a function of electrode configuration. Thus patients with strong susceptibility to channel interaction will produce higher speech recognition scores for sequential as opposed to simultaneous speech strategies. Speech strategy performance was evaluated with acute listening trials where patients identified consonants and vowels with each of the following speech strategies, listed in order from sequential stimulation to fully simultaneous stimulation: Continuous Interleaved Sampler (CIS), Paired Pulsatile Sampler (PPS), Hybrid Analog Pulsatile (HAP), and Simultaneous Analog Stimulation (SAS). Preliminary results showed that a greater degree of channel interaction was associated with the highest speech recognition scores for CIS, midrange scores for HAP, and the lowest scores for SAS. [Work supported by NIH.]

4aSCa5. Cochlear-implant effects on sibilants. Melanie Matthies (Dept. of Commun. Disord., Boston Univ., 635 Commonwealth Ave., Boston, MA 02215, matthies@bu.edu), Joseph Perkell, Jennell Vick, and Majid Zandipour (MIT, Cambridge, MA 02139)

To investigate sibilant production in persons with cochlear implants, this study compared acoustic analyses of three sets of speech stimuli: (1) A synthesized continuum of sibilants from ''sa'' to ''sha'' was constructed where the characteristics of the sibilant noise as well as the F2 and *F*3 transitions were varied. (2) The synthesized speech was processed through a cochlear-implant simulation program [cf. Shannon *et al.*, Science **270**, 303–304 (1995)]. (3) Longitudinally recorded speech production data were obtained from postlingually deafened adults who had received cochlear implants. Formant frequencies and amplitudes were obtained at midvowel and at vowel onset. The amplitude in the sibilant at the frequency of vowel *F*3 was measured, as were the spectral characteristics of the sibilant noise for all three sets of speech. Judgments about the identity and quality of the synthesized sibilants were obtained from a panel of listeners with normal hearing. Correlations indicated that the selected acoustic cues were highly salient to the listeners for both the unprocessed and cochlear-implant-simulated sibilants. Observed improvements in sibilant production from pre-CI to post-CI are likely due to the improved availability of these salient cues for auditory feedback and self-monitoring. [Work supported by MURI Z883402, NIDCD DC03007.]

4aSCa6. Feature extraction in human speech recognition. J. B. Allen, M. Furst, L. Saul, and M. Rahim (AT&T Labs, Florham Park, NJ 07932)

In 1955 George Miller and Patricia Nicely explored the human speech code and discovered that a five-channel feature code is used to extract phonemes from speech. Their mutual information analysis of the confusion matrices of filtered noisy speech showed that the five channels were virtually independent of each other, accounting for the robustness of speech communication quantified by Miller, Heise, and Lichten in 1951. While the 1955 analysis was a major step forward, it did not lead to practical results, since there are no known methods for extracting these features from the speech. This weakness is the basis of the present research program, to define the signal processing required to robustly extract features from narrow bands of speech. Our emphasis is to work in narrow bands, starting from critical bands. In this paper we discuss the first step of this process, which is to measure (a) the jnd for the detection of speech in noise and (b) the jnd for residue pitch given narrow bands of speech in noise. We have been investigating these features with the ultimate goal of trying to build a speech recognition machine based on distinctive feature signal processing.

4aSCa7. Vowel perception in varied symmetric CVC contexts. Jim Talley (Dept. of Linguist., Univ. of Texas, Austin, TX 78712)

In the three-decade-long debate over static versus dynamic specification of vowels, perceptual studies in which subjects are tasked with identifying naturally spoken vowels under various ablation conditions have been a mainstay. While not directly producing an understanding of how humans go about recognizing this major subclass of phones, this type of study [e.g., Strange, Jenkins, and Johnson, J. Acoust. Soc. Am. **74**, 695–705 (1983)] has provided compelling results which must be accounted for in any successful theory of vowel perception. This paper presents results from yet another perceptual study of human vowel identification under ablation conditions. This study uses CVC syllables spoken rapidly by three male speakers in a carrier sentence. Syllables consist of ten American English vowels in each of four consonantal contexts (b_b, d_d, g_g, and h_d). The conditions studied are silent centers (SC), centers only (CO), and the control condition (full). A very robust hierarchy of full $>$CO$>$SC is found. Consonantal contexts also have a clear ordering (h_d$>$b_b$>$d_d$>$g_g) with respect to the ease with which they are perceived. Interesting interactions between vowels and their contexts are also evident.

4aSCa8. Vowel formant discrimination in natural speech. Chang Liu and Diane Kewley-Port (Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405, liuchang@indiana.edu)

Vowel formant discrimination in syllables, phrases, and sentences was measured for nearly natural speech resynthesized by STRAIGHT [Kawahara *et al.*, Speech Commun. 187–207 (1999)]. The purpose of experiment 1 was to validate the use of STRAIGHT to modify formants by comparing vowel discrimination in formant-synthesized and -resynthesized isolated vowels. The resulting DLs for *F*1 and *F*2, calculated in barks, were the same as previously reported. In experiment 2, DLs were measured in natural speech for *F*1 and *F*2 for the vowels /ih, eh, uh, ae/. Different phonetic contexts included syllables, phrases, and sentences, as well as sentences with the addition of an identification task. Results showed that longer phonetic context was not a strong factor in degrading DLs, although performance for syllables was significantly better than for sentences. The addition of the identification task did not have a significant effect on vowel discrimination. In comparison to the formant-synthesized speech in Kewley-Port and Zheng [J. Acoust. Soc. Am. **106**, 2945–2958 (1999)], vowel discrimination was similar for syllables, but significantly degraded for more complex contexts. However, the magnitude of the results suggests that vowel formant discrimination is similar for synthetic and natural speech in a variety of phonetic contexts. [Work supported by NIHDCD-02229.]

4aSCa9. Extracting vowel characteristics from smoothed spectra. Jean-Sylvain Lienard (LIMSI-CNRS, B.P.133, 91403 Orsay Cedex, France, lienard@limsi.fr) and Maria-Gabriella Di Benedetto (Universita' di Roma La Sapienza, 00184 Rome, Italy)

The object of this study was to define vowel spectral features by means of a spectral representation making no specific reference to formants and fundamental frequency, though preserving vowel identity and extralinguistic factors such as effort and gender. This investigation complements the analysis in Lienard and Di Benedetto [J. Acoust. Soc. Am. **106**, 411–422 (1999)] which used manually measured formants of the same data. The database consisted of isolated French vowels, uttered by male and female speakers, with different vocal efforts. The analysis was carried out by selecting one segment along the vowel corresponding to a joint maximum of stability and level. A spectral envelope, based on prominent spectral peaks selected by a frequency-domain masking function, was computed. The frequency axis was Bark-transformed. The envelope was then decomposed by successive smoothings (cepstral filtering). The output was a representation where vowels are characterized by the presence or absence of energy concentration at different frequency locations (''spectral bumps''). Spectral bumps, although related to formants, do not correspond to formants themselves; rather, they reflect an integrated information similar to what might be used by the perceptual system.

4aSCa10. Effect of masking of spectrally reduced speech with modulated noise. Angélique Grosgeorges, Frédéric Berthommier (Institut de la Commun. Parlée, 46, Av. F. Viallet, 38031 Grenoble Cedex, France), and Christian Lorenzi (LPE, 92774 Boulogne-Billancourt, France)

Accurate recognition performance can be obtained with spectrally degraded speech [R. V. Shannon *et al.*, Science **270**, 303–304 (1995)]. Part of this experiment was initially reproduced for 16 VCVCV stimuli (a/C/a/C/a) spectrally degraded and presented in quiet. Then, the contribution of low- and high-frequency amplitude modulations (AM) of speech and noise masker presented at 6-dB SNR was investigated, using a 2 factors/2 levels experimental design: The speech envelope was low-pass filtered at 10 or 500 Hz; the masker was a white noise with low-frequency AM components (<8 Hz), which was sinusoidally amplitude modulated at 100 Hz or not. In addition, the low-frequency AM components of the masker were co-modulated across the four frequency bands or not. Compared to the scores obtained in quiet, a decrease of about 50%–75% in the transmitted information for place and voicing cues was observed in all conditions, and transmission of manner was dramatically suppressed. As expected, recognition of spectrally degraded speech was strongly affected by temporal masking, but the detrimental effects of noise were not uniform for all speech features, giving some insight about their respective temporal encoding. [Work supported by RESPITE.]

4aSCa11. Gross versus detailed spectral cues in spectrally distorted speech. Michael Kiefte[a] (Dept. of Linguist., Univ. of Alberta, Edmonton, AB T6G 2E7, Canada)

An experiment was designed to address the importance of gross versus detailed spectral features in the perception of stop-consonant place of articulation. This is done by presenting listeners with noise-vocoded stimuli in which the spectral resolution of the processed tokens is severely reduced. While it can be assumed that such a transformation will effectively eliminate detailed spectral features, such as burst-peak and formant frequencies, previous work has shown that this is not necessarily the case [M. Kiefte, J. Acoust. Soc. Am. **106**, 2273 (1999)]. Nevertheless, it is shown that gross spectral features, such as spectral tilt and compactness, are better able to predict listeners' identifications of noise-vocoded stimuli. While this suggests that listeners attend primarily to gross spectral shape cues, which are assumed to be particularly robust in such distorted speech, it is also shown that detailed spectral features are better able to model listeners' responses to *undistorted* speech. [Work supported by SSHRC.]
[a] Now at the Dept. of Psychology, University of Wisconsin, Madison, WI.

4aSCa12. Long-term speech spectra for Taiwanese–Mandarin bilingual speakers. Chin-Hsing Tseng, Hsiao-Chuan Chen (Dept. of Special Education, Natl. Kaohsiung Normal Univ., Kaohsiung, Taiwan, ROC), and Shing Chen (Industrial Technol. Res. Inst., Hsin-Chu, Taiwan, ROC)

Audiologists and otologists in Taiwan often rely on the English normative data during the hearing aid gain prescription procedure. Thus, the measurement of long-term speech spectra for both Taiwanese and Mandarin was deemed a research priority. Average long-term rms 1/3-octave band speech spectra were generated for 60 Taiwanese–Mandarin speakers with gender, age, and language as the three major independent variables. The male speakers differed significantly from the females in the low-frequency bands across languages and age groups. However, a slight gender difference in the 4000-Hz band was observed only in the young adult group (aged 20–40) for both Taiwanese and Mandarin. Across languages, the elder group (aged 41–60) differed significantly in the low-frequency bands for both genders, but high-frequency differences existed only in the females. The overall spectra will provide a useful reference for hearing aid gain prescription in the Chinese-speaking communities. Cross-language comparisons can also shed insight on the speech communication processes. [Work supported by the Council of Labor Affairs.]

4aSCa13. Effects of spectral mismatch on sentence recognition for noise-band processors. John J. GalvinIII and Qian-Jie Fu (Dept. of Auditory Implants and Percept., House Ear Inst., 2100 W. 3rd St., Los Angeles, CA 90057)

For cochlear implant (CI) users, the cochlear insertion depth and limited spatial extent of the electrode array causes a mismatch between the frequency content of speech signals and the place of stimulation along the cochlea. To examine the effects of such a spectral mismatch on HINT sentence recognition in noise, five normal-hearing (NH) listeners were tested using noise-band speech processors designed to simulate a variety of cochlear insertion depths. The speech reception threshold (SRT) was measured for two conditions: (1) analysis and carrier bands matched to the insertion depth, and (2) analysis bands fixed over a wide range while carrier bands varied according to insertion depth. In the matched condition, subjects' performance remained flat along the entire range of insertion depths; on average, the SRT was 5 dB lower than the mismatched condition, indicating that a frequency-limited, but spectrally matched processor provided more benefit. In the mismatched condition, performance was flat over most of the range of insertion depths and declined only for extremely basal or apical insertions. The results indicate that electrode insertion depth and the resulting spectral mismatch is probably not the limiting factor affecting speech recognition or contributing to the variability in performance observed among CI users.

4aSCa14. Effects of lexical and talker variability factors on spoken word recognition in native and non-native, normal and hearing-impaired listeners. Sumiko Takayanagi, Donald D. Dirks, Anahita Moshfegh, P. Douglas Noffsinger (UCLA School of Medicine, Los Angeles, CA 90095, takayana@ucla.edu; Natl. Ctr. for Rehab. Auditory Res., OR 97207, and West LA VAMC, Los Angeles, CA 90073), and Stephen A. Fausti (Natl. Ctr. for Rehab. Auditory Res., Portland, OR 97207)

The effects of lexical difficulty and talker variability on word recognition were examined in four groups of listeners: native English/normal hearing; native English/hearing impaired; non-native English/normal hearing; and non-native English/hearing impaired (hearing level matched to the native hearing impaired). Lexical difficulty was measured by the difference in performance to 75 lexically ''easy'' and ''hard'' words based on word frequency and Neighborhood Activation Theory [Luce and Pisoni (1998)]. The effect of talker variability was measured by the difference in performance between single and multiple talker (nine talkers) conditions. The familiarity of the 150 words was rated on a seven-point scale. An up–down adaptive procedure was used to determine the sound pressure

level for 50% performance. Non-native listeners in both normal and hearing-impaired groups required a greater intensity for equal intelligibility than for the comparative native normal and hearing-impaired listeners. Results, however, showed significant effects of lexical difficulty and talker variability in all four groups. Structural equation modeling demonstrated that an auditory factor estimated by pure tone average, etc., accounts for four times more variance to performance than does a linguistic fluency factor measured by word familiarity ratings and native versus non-native status, however, the linguistic fluency factor is also essential to the model fit.

4aSCa15. Speaking clearly for learning-disabled children: Sentence perception in noise. Ann R. Bradlow (Dept. of Linguist., Northwestern Univ., 2016 Sheridan Rd., Evanston, IL 60208), Nina Kraus, and Erin Hayes (Northwestern Univ., Evanston, IL 60208)

This study sought to establish connections between two findings regarding speech perception in special populations. First, children with learning problems often have difficulty perceiving speech in noise. Second, naturally produced clear speech is more intelligible for adult hearing-impaired listeners. Given recent interest in the use of digital enhancement techniques in speech and language training procedures for learning-disabled children, we wanted to investigate whether naturally produced clear speech yields perception benefits for this population. We presented children with learning problems ($n=46$) and normal children ($n=17$) with sentences embedded in noise. Factors that varied were speaking style (plain versus clear) and signal-to-noise ratio (high versus low). Results indicated that: (1) children with learning problems had poorer sentence-in-noise perception than normal children; (2) both groups had poorer perception with increasing noise levels; and (3) both groups benefited from naturally produced clear speech. There were no significant two- or three-way interactions between listener group, noise level, and speaking style. In combination with acoustic analyses of the plain and clear sentences, these findings have broad implications for our understanding of the speech perception deficits in this population, and the design of appropriate training procedures. [Work supported by NIH-NIDCD Grants DC 03762 and DC 01510.]

4aSCa16. Adaptive thresholds for spoken sentence materials. Theodore Bell (Dept. of Commun. Disord., King Hall B119, 5151 State University Dr., Cal State L.A., Los Angeles, CA 90032-8170), Richard Wilson (Quillien Veteran Affairs Medical Ctr., Mountain Home, TN), Jane Enrietto, and Carmen Morris (Cal State L.A., Los Angeles, CA 90032-8170)

Two experiments document the reliability of a speech discrimination test designed for use with an adaptive procedure to measure thresholds for sentences and words. Stimulus words were categorized as either high usage or low usage based on published indices of word usage, and also as either high confusability or low confusability depending on phonetic similarity to other words in the lexicon. The first experiment employed 20 young normal-hearing adults, each presented with two 20-item lists of sentences from each of the four lexical categories. One week later, the testing sequence was repeated. A five-factor ANOVA was employed using word usage, word confusability, list sample, and test–retest delay; an adaptive procedure, converging determined sentence thresholds in dB level. Test–retest thresholds, both immediate and delayed, were within 2 dB for all lists. There was a 1.5-dB difference between high- and low-usage words. A second experiment examined thresholds for groups of hearing-impaired individuals categorized by audiometric configuration and speech discrimination scores. Each subject group was presented two sentence lists varying in lexical characteristics based on the Neighborhood Activation Model. The difference between lexical items varied significantly with speech discrimination scores and audiometric configuration of the listeners. [Work supported by Veteran Affairs R&D Contract V621P-3664.]

4aSCa17. Co-modulation masking release (CMR) in consonant recognition. Bom Jun Kwon (Dept. of Auditory Implant and Percept., House Ear Inst., 2100 W. 3rd St., Los Angeles, CA 90057-1922)

In the psychoacoustic phenomenon of co-modulation masking release (CMR), detection of a target in a narrow-band noise is improved in the presence of a co-modulated flanking noise band. The present study investigates whether CMR could improve speech perception. Consonant recognition scores of bandlimited speech, 1200–2200 Hz, were measured under a modulated noise of the same bandwidth, and a second band of noise (5000–7000 Hz), which was either co-modulated or non-co-modulated with the first band. With a hypothesis that CMR improves discrimination of phonemes that are characterized by temporal cues, i.e., amplitude envelope, rather than spectral cues, two groups of consonants were selected for testing: stop consonants (similar temporal envelopes, but different spectra) and alveolar–dental consonants (similar spectra, but different temporal envelopes). Results agreed with the hypothesis: significantly higher recognition scores were obtained with the co-modulated noise band for alveolar–dental consonants, while the difference was marginal for stop consonants. This indicates that CMR can improve speech perception as well as psychophysical thresholds. Co-modulation in the flanking band seems to enable listeners to better ''hear out'' subtle temporal speech cues in noise, leading to better recognition. [Work supported by Quota scholarship and NIDCD.]

4aSCa18. Effects of filtering preceding noise on the intelligibility of voiced plosives. William A. Ainsworth (MacKay Inst. of Commun. and Neurosci., Keele Univ., Staffordshire ST5 5BG, UK) and Teresa Cervera (Univ. of Valencia, Spain)

The intelligibility of voiced plosives in continuous noise is greater than in noise which is coincident with the speech [W. A. Ainsworth and G. Meyer, J. Acoust. Soc. Am. **96**, 687–694 (1994)]. It is suggested that this could be due to some adaptation process. In order to test this hypothesis, experiments have been performed to measure the intelligibility of voiced plosives in isolated syllables with different durations of preceding noise. The effects of the spectral shape of the noise have also been examined. It was found that, on average, 800 ms of band-stop filtered noise preceding the syllable increased the recognition rate of the plosives by 14%, band-pass filtered noise by 12%, low-pass filtered noise by 7%, but high-pass filtered noise by only 3% compared with co-gated noise. The time constant of the adaptation process was estimated to be 165 ms. [Work supported in part by EU TMR Network SPHEAR.]

4aSCa19. Across-frequency timing in speech perception: Effective processing of contrasting temporal speech patterns requires proper synchronous alignment. Eric W. Healy and Sid P. Bacon (Psychoacoustics Lab., Dept. of Speech and Hearing Sci., Arizona State Univ., P.O. Box 871908, Tempe, AZ 85287-1908, ewh@asu.edu)

Although narrow-band temporal speech patterns produce virtually no intelligibility when presented individually, a contrasting pair of such patterns can produce appreciable intelligibility [E. W. Healy, Dissertation, University of Wisconsin–Milwaukee (1998)]. In the current study, the role of across-frequency timing in speech perception was investigated using these contrasting temporal speech patterns which allow the strict isolation of across-frequency integration. Across-band timing was investigated by varying the synchrony of paired temporal patterns. Individual temporal patterns were created by amplitude modulating tones at 750 and 3000 Hz to follow the temporal fluctuations of corresponding 1/3-octave sentences. The pair of individually unintelligible narrow-band temporal patterns were mixed and presented to listeners either in temporal synchrony, or with offsets of 12.5, 25, 50, or 100 ms. High intelligibility scores in the aligned condition were reduced when as little as 12.5 ms of misalignment was introduced, and performance fell to values near zero at the largest misalignment. These results indicate that normal-hearing young listeners are relying heavily on across-channel timing information provided by the syn-

chronous temporal patterns. Further work employing this technique may provide insight into the role of timing in the suprathreshold speech processing deficits of elderly or hearing-impaired listeners. [Work supported by NIDCD.]

4aSCa20. Discriminant analysis of acoustic cues to three Green Mong tones. Jean E. Andruski (ASLP, 581 Manoogian Hall, Wayne State Univ., Detroit, MI 48202)

Green Mong is a tone language with an inventory of seven contrastive tones. The tonal system of Green Mong incorporates both pitch and phonation-type distinctions, and appears to make use of many of the acoustic parameters which have been shown to correlate with phonation-type cross-linguistically. This study looks at three Green Mong tones that have similar pitch contours, but are distinguished by the distinctive use of breathy, creaky, and modal phonation. Acoustically, the three tones are distinguished to some degree on the basis of pitch, relative amplitude of lower and higher harmonics (H1–H2), duration, vowel quality, and voice onset time. Discriminant analysis is used to estimate the value of these different cues for classifying tokens into tone categories based on observed characteristics in tone production. Discriminant functions were generated using measurements from 270 CV words produced by six speakers, for which every combination of three vowels (/i/, /u/, /a/) five consonants, and three tones (breathy, creaky, modal) was attested. Results of the analyses will be presented, as well as results for classification of a separate, test set of 102 CV tokens produced by the same six speakers.

4aSCa21. Relation of pitch glide perception and Mandarin tone identification. Terry L. Gottfried (Dept. of Psychol., Lawrence Univ., Appleton, WI 54912) and Devon Riester (Kenyon College, Gambier, OH 43022)

Second language (L2) learners differ considerably in their ability to learn new phonetic contrasts in the L2. For example, Gottfried and Suiter (1997) observed that many native speakers of American English were highly accurate in their identification of Mandarin Chinese tones, while others, with similar years of experience in learning Mandarin, were very poor. The current study compared listeners' identification of Mandarin tones with their ability to perceive the pitch direction of sine-wave glides. Sine waves of 400 ms with starting frequencies of 250 and 300 Hz were created with 0- to 50-Hz sweeps up or down from the initial frequencies. Listeners who had no experience with Mandarin indicated whether the pitch went up, down, or remained the same. After testing on the sine waves, listeners identified the tone (high-level, mid-rising, low-dipping, and high-falling) of intact and silent-center (with only the initial and final portions) Mandarin /l/-vowel syllables. Analysis revealed that listeners with musical training performed significantly better on intact Mandarin syllables than nonmusic majors, but no difference was observed for silent-centers. Voice majors seemed to have a particular advantage. Significant correlations were also found between tests of glide perception and tone identification (ranging from 0.530 to 0.642).

4aSCa22. Effects of time reversal on consonant identification. Sneha V. Bharadwaj and Peter F. Assmann (School of Human Development, Univ. of Texas at Dallas, Box 830688, Richardson, TX 75083-0688, snehab@utdallas.edu, assmann@utdallas.edu)

Time-reversed speech has the same long-term amplitude spectrum as natural speech, but is largely unintelligible. To study the perceptual effects of disrupting the temporal integrity of formant transitions on consonant identification, CVC syllables were divided into brief, successive, nonoverlapping segments and each segment was time reversed. In experiment 1, CVC syllables from two talkers were time reversed using 20, 40, 60, 80, or 100-ms temporal windows. The initial consonant was either /p/, /t/, /k/, /b/, /d/, /g/, /s/, /ʃ/, /f/, or /v/. The medial vowel was /i/, /u/, or /ʌ/, and the final consonant was /t/. Identification accuracy (12 listeners) showed a progressive decline from 99% for unmodified syllables to 63% for the 100-ms window. Time reversal was less disruptive when syllables were presented in a natural carrier sentence, dropping to 75% for a 100-ms time window. Experiment 2 replicated this pattern of findings with a larger sample of ten talkers. The effects of time reversal were not uniform across the 12 consonants, but showed highly disruptive effects on the identification of stops and relatively minor changes in fricatives. Overall, the results suggest that time reversal provides a useful method for quantifying the role of dynamic information in consonant perception.

4aSCa23. Analysis-by-synthesis and acoustic landmarks. Richard S. McGowan (CReSS LLC, 1 Seaborn Pl., Lexington, MA 02420)

In the 1950s and 1960s Ken Stevens and Morris Halle proposed a model for human speech recognition that incorporated analysis-by-synthesis. In the more recent past the notion of acoustic landmarks has arisen from the speech group at MIT as part of a model for human lexical access from features by human listeners. Here, it is proposed that these ideas are fruitful ones for incorporation into a method (constraint and refinement) for training a machine with its own vocal tract to imitate human speech, while learning the relation between articulation and acoustic output. The machine would be trained in stages starting with the attempts to imitate the sequence of relatively open articulations and closed articulations. Landmarks, based on acoustic criteria, would be placed near the peaks in the amplitude envelope, allowing the machine to constrain the type of articulatory activations. This constraining step would involve, initially, an overly restrictive set of phonotatic rules, such as all sequences are obstruent-vowel sequences. The refining step would be the analysis-by-synthesis with the machine's vocal tract. As the stages progress, the machine would attend to more acoustic properties of the speech signal with different kinds of landmarks and increase the generality of the phonotatic rules.

4aSCa24. Auditory robustness and duration of vocalic cues. John J. Ohala and Alan C. Yu (Dept. of Linguist., 1203 Dwinelle Hall, Univ. of California, Berkeley, Berkeley, CA 94720-2650, ohala@socrates.berkeley.edu)

Building on an idea of Stevens [ICPhS (1980)] and Stevens and Keyser [Language (1989)], Lang and Ohala (LO) [ICSLP (1996)] hypothesized that the most commonly used features to differentiate vowels in languages would be those few that can be detected in a relatively short time. Other features could be used to augment the vowel inventory at the expense of requiring more time for their perception. They found support for this idea in a perceptual study of end-truncated English vowels. We conducted a similar experiment on the perception of 18 contrasting vowels in Cantonese including monophthongs and diphthongs and both rounded and unrounded front vowels. The stimuli were vowel fragments starting at 30 ms after an initial [s] and end truncated into pink noise at 20-ms intervals thereafter. Twenty Cantonese heard randomized sets of these stimuli and attempted to associate each with the vowel in 18 different key words symbolized by Chinese characters. The resulting confusion matrices of the shorter fragments revealed not only the expected high confusions between monophthongs and diphthongs but also between rounded and unrounded front vowels. The latter contrast, which is indeed used less in language's vowel inventories, must require more time to perceive, supporting the LO claim.

Session 4aSCb

Speech Communication and Archives and History: Speech Communication Within the Acoustical Society of America

Patricia A. Keating, Chair

Department of Linguistics, University of California, Los Angeles, 3125 Campbell Hall, Los Angeles, California 90095-1543

Chair's Introduction—11:10

11:15

4aSCb1. Speech communication within the Acoustical Society of America. Katherine Harris (Haskins Labs., 270 Crown St., New Haven, CT 06511), Peter Ladefoged (UCLA, Los Angeles, CA 90095), and Kenneth Stevens (MIT, Cambridge, MA 02139)

Speech communication has been an important field of the Acoustical Society from its earliest years. Harvey Fletcher, the first President of the Society, along with many colleagues at Bell Telephone Laboratories, provided new data and techniques for investigating speech. The invention of the sound spectrograph during World War II led to greater understanding of the physics of speech, and this in turn produced developments in speech synthesis, summarized through the mid 1980's by Klatt [J. Acoust. Soc. Am. **82**, 737–793 (1987)]. Physiological, psychological and linguistic aspects of speech have also been much investigated within the Society.

A NOTE ABOUT THE ASA HISTORY LECTURE SERIES

In 1997, the ASA Committee on Archives and History conceived a plan for a series of invited lectures on each of the technical areas of the Society which would memorialize the significant achievements and milestones of each of its twelve technical committees and one interdisciplinary technical group during the first three quarters of the Society's first century.

With the cooperation of the technical committees, distinguished individuals are selected to review the history of their particular technical speciality and present a lecture which shows how that activity has developed and has contributed to the Society at large and to the broad field of acoustics as well. At the meetings in Columbus, Ohio and Atlanta, Georgia, the first four History Lectures were presented: Gabriel Weinreich on Musical Acoustics, Robert Beyer and David Blackstock on Physical Acoustics, Stanley Ehrlich on Engineering Acoustics and Ewart Wetherill on Architectural Acoustics. At each subsequent meeting two additional lectures will be scheduled including those in Noise and Speech Communication at this meeting in Newport Beach.

The invited lecturers have been asked to prepare a written manuscript of their lectures which will be published in a commemorative book for the 75th Anniversary of the Society to be celebrated in 2004. The Archives and History Committee and the individual technical committees/group welcome comments and suggestions on both the History Lecture Series and on the proposed ASA Diamond Anniversary Book. Volunteers to assist the committees would be most welcome too. Contact Henry Bass, Chair, Committee on Archives and History, pabass@sunset.backbone.olemiss.edu

Session 4aUW

Underwater Acoustics and Signal Processing in Acoustics: Acoustic Time Reversal and Acoustic Communications

David R. Dowling, Chair
Department of Mechanical Engineering and Applied Mechanics, University of Michigan, 2019 Autolab 2121, Ann Arbor, Michigan 48109-2121

Chair's Introduction—8:10

Contributed Papers

8:15

4aUW1. Time-reversal pulse compression in shallow water with a single-element source–receiver. Charles F. Gaumond and David Fromm (Acoust. Div. Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375-5320)

A time-reversing, full-water-column, source–receiver array has been previously demonstrated to focus and time compress a field at the position of the source through the use of all the propagating modes of the water column [W. S. Hodgkiss *et al.*, J. Acoust. Soc. Am. **105**, 1597–1604 (1999)]. Simulations are presented that show the limitations of using a single-element echo repeater instead of a full-water-column array. The system consists of a source located at the center of a vertical line array (VLA) and a distant, single-element echo repeater (SEER). The echo field produced by the SEER has very limited vertical extent of time compression at the range of the source and VLA. This vertical extent can be extended with the use of received signals from a varying number of adjacent hydrophones. The tradeoffs involved with extending the compression are shown. The simulations are generated using RAM in a simple downward-refracting shallow-water environment. [Work supported by ONR.]

8:30

4aUW2. Broadband performance of a time-reversing array in shallow water. Karim G. Sabra and David R. Dowling (Dept. of Mech. Eng., Univ. of Michigan, Ann Arbor, Ann Arbor, MI 48109-2121, drd@engin.umich.edu)

Future underwater communication systems may be developed from the spatial and temporal focusing properties of a time-reversing array (TRA). These properties ensue from the interactions between the underwater environment and the TRA. In a shallow-water sound channel, the influence of the TRA element number (6, 12, 20) and spacing (5 m, 3 m), and source-array range (1 to 40 km) on the spatial focusing and temporal compression will be presented. Here, the results show that compromises are necessary. Short ranges produce better spatial focusing while longer ranges produce better temporal compression. In addition, theoretical results for the performance of the TRA in an underwater environment with an ambient noise field has been evaluated for a family of pulses that span the range from narrow- to broadband. The theory, which has been the subject of previous presentations, accounts for the number of array elements and the time–bandwidth product of the signal. Based on this theory and numerical experiments, we will first investigate the effect of the sound channel characteristics on the performance of broadband signals and then compare broadband signals to narrow-band signals for several source-array ranges. [Work sponsored by the Office of Naval Research.]

8:45

4aUW3. Robust diagnosis for the focal stability in time-reversal acoustics. H. C. Song, W. A. Kuperman (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0238), T. Akal (SACLANT Undersea Res. Ctr., 19138 La Spezia, Italy), W. S. Hodgkiss, S. Kim, and G. Edelmann (Univ. of California, San Diego, La Jolla, CA 92093-0238)

For a time-independent medium, one can use stored probe pulses to focus on a specific location in time-reversal acoustics. However, the temporal variability of the ocean is expected to limit such a procedure. In the earlier low-frequency (445 Hz) experiments, we found that probe pulses up to 1 week old still produced a significant focus at the original probe source location. In the recent high-frequency (3500 Hz) experiments, however, stability of the focus was limited to less than an hour or even a few minutes, depending upon the oceanographic variability since the acoustic wavelength is much smaller than the one for 445 Hz. A quick and robust diagnosis for the focal stability appears to be the modal group velocity versus phase velocity (gp curves) derived from the sound-speed profile, where the phase speed is directly related to the mode number and the slopes of the curves separate different groups of modes. In this paper, we investigate the stability and scintillation of the focus observed during the May 2000 time-reversal experiment using the gp curves.

9:00

4aUW4. Spatial resolution of time reversal arrays in shallow water. Seongil Kim, Geoffrey F. Edelmann, William S. Hodgkiss, William A. Kuperman, Hee Chun Song (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0238), and Tuncay Akal (SACLANT Undersea Res. Ctr., 19138 La Spezia, Italy)

A series of time reversal (phase conjugation) experiments was performed in shallow water including a range-dependent slope environment. Time reversal arrays implemented with center frequencies of 445 and 3500 Hz achieved sharp focal regions up to ranges of 30 and 13 km, respectively, in 110–130 m shallow water. Analysis with an image method indicates that the focal size approaches the diffraction limit of an array for given waveguide conditions, i.e., waveguide geometries and attenuation. The measured focal size has implications for the maximum achievable resolution of matched field processing, which is a computational implementation of the phase conjugation process. [Work supported by ONR.]

9:15

4aUW5. Steering nulls in TRM with distortionless response at focal location—Theory and simulation. J. S. Kim, W. A. Kuperman, and H. C. Song (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093, jkim@mpl.ucsd.edu)

The time-reversal mirror uses the received signal from a probe source to refocus the signal at the probe source location by back-propagating the time-reversed signal. In this study, a method to place a null at an arbitrary position while maintaining a distortionless response at the probe source location, assuming that the transfer function is known at that point, is proposed and demonstrated via simulation in Pekeris ocean waveguide. It is further shown that the null can be placed in the vicinity of probe source at the same depth using the theory on the invariants of ocean waveguide, without any environmental information. The approach is based on the adaptive weighting, of which formulation is a classical minimization problem with a constraint known as the minimum variance with distortionless response. The proper weighting for an array is applied before backpropagation. Finally, it is demonstrated that the null can be placed at the probe source location.

9:30

4aUW6. Applying the time-reversal operator to detection in a shallow-water environment. David Fromm and Charles F. Gaumond (Acoust. Div. Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375-5320)

The decomposition of the time-reversal operator (DORT) is a method of isolating scatterers with a multiple-source/multiple-receiver system [C. Prada *et al.*, J. Acoust. Soc. Am. **99**, 2067–2076 (1996)]. A series of simulations is shown that elucidates the resolution of a DORT system in a reverberant, shallow-water environment. The system consists of vertical line arrays of sources and receivers in monostatic and vertically bistatic configurations. Two basic resolution problems are considered: the ability to discriminate two scatterers in the water column and the ability to detect a single scatterer in the presence of bottom clutter. The performance of the DORT system is shown relative to array configuration, location of scatterer in the water column, and the level and type of bottom clutter. [Work supported by ONR.]

9:45

4aUW7. Forward-scatter barrier with a time-reversal mirror. W. A. Kuperman (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0701), T. Akal (SACLANT Undersea Res. Ctr., 19138 La Spezia, Italy), W. S. Hodgkiss, S. Kim, G. Edelmann, and H. C. Song (Univ. of California, San Diego, La Jolla, CA 92093-0238)

The classic difficulty in constructing an acoustic trip line barrier is that the forward-scattered field from the target must be extracted from the (usually) much more intense direct blasting arriving beam, i.e., ''looking into the sunlight effect.'' During the time-reversal experiment conducted in May 2000, we investigated the forward-scatter barrier concept using a six-element transponder at 65-m depth drifting along with a ship, which traversed the trip line between two moored vertical arrays separated by 5 km in 110-m water depth. A 10-ms cw pulse from a probe source (PS) located at the bottom of the vertical receive array (VRA) is received and time-reversed at the source/receive array (SRA). The transmitted time-reversed signal is then refocused at the PS location of the VRA. This time-reversed signal is also captured by the transponder simulating the forward-scatter target and is retransmitted to the VRA with various amplitudes simulating different forward-scatter target strengths. With an appropriate time delay, the main blast and the transponder signal arrived at the VRA simultaneously. For comparison, one-way broadside transmissions were also made.

10:00–10:15 Break

10:15

4aUW8. Underwater acoustic communication by passive phase conjugation: Theory and experiment. Darrell R. Jackson, Daniel Rouseff, Warren L. J. Fox, Christopher D. Jones (Appl. Phys. Lab., College of Ocean and Fishery Sci., Univ. of Washington, Seattle, WA 98105, rouseff@apl.washington.edu), James A. Ritcey (Univ. of Washington, Seattle, WA 98105), and David R. Dowling (Univ. of Michigan, Ann Arbor, MI 48109)

A new method for coherent underwater communication called passive phase conjugation is evaluated. The technique takes its name because of conceptual similarities to active phase conjugation methods that have been demonstrated in the ocean [Kuperman *et al.,* J. Acoust. Soc. Am. **103**, 25–40 (1998)]. In contrast to active techniques, however, the array in passive phase conjugation need only receive. This makes the method plausible for scenarios where spatially compact sources might be communicating to a distant receive-only array. Compared to other approaches for coherent communication, the computational burden is low allowing the method to be evaluated in the field in nearly real-time. Results from an experiment conducted in Puget Sound near Seattle in May 2000 are reported. Various modulation schemes and array geometries were employed. Measurements were made at several ranges and water depths in a range-dependent environment. [Work supported by ONR.]

10:30

4aUW9. Self-equalization communications using a time-reversal mirror. G. Edelmann (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0238), T. Akal (SACLANT Undersea Res. Ctr., 19138 La Spezia, Italy), W. S. Hodgkiss, S. Kim, W. A. Kuperman, and H. C. Song (Univ. of California, San Diego, La Jolla, CA 92093-0701)

A time-reversal mirror (TRM) was used to transmit self-equalization communication sequences in both range-independent and range-dependent environments. The TRM consisted of 29 source/receive elements centered at 3.5 kHz. A 2-ms pure-tone pulse was transmitted from a probe source at a range of 10 km from the TRM. A vertical receive array with 32 elements was colocated with the probe source to measure the spatial resolution of the TRM focus and bit error rate. The communication sequences were encoded using BPSK and higher-order phase-shift keying. As a comparison, single-source and broadside communication transmissions were also made.

10:45

4aUW10. Using spatial coherence to improve the performance of underwater acoustic communications. T. C. Yang (Naval Res. Lab., Washington, DC 20375) and Martin Siderius (SACLANTCEN, La Spazia, Italy)

Phase-shifted keying modulation schemes offer an efficient use of signal power for a given bandwidth in a communication channel. Successful communication often requires multichannel data (spatial diversity) to combat the signal fading due to random fluctuations in the channel. For each channel, a signal-to-noise ratio (SNR) of >10 dB is often required to estimate the channel impulse response for proper placement of the tap coefficients. Array beamforming can be used to improve the SNR, but when to use spatial diversity or beamforming is an issue long debated but not yet satisfactorily answered. It is shown with at-sea data that an effective use of beamforming versus spatial diversity depends on the phone spacing relative to the signal coherence. Beamforming requires the signals to be correlated between the sensors, whereas spatial diversity requires the opposite. Neither can fulfill the job of the other. By choosing phones according to spatial coherence and combining beamforming with diversity, bit error rates of the order of ~ 0.001 were demonstrated using Advent 99 data over a period of $\sim$ 9 hours with an SNR as low as ~ 1 dB at the phone level. [This work is supported by the Office of Naval Research and SACLANTCEN.]

11:00

4aUW11. Analytic performance evaluation of coherent equalizers in time-varying acoustic channels. James C. Preisig (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Coherent equalization is one of the primary methods of implementing high-speed acoustic communications systems in underwater environments. This environment is often characterized by extended delay and Doppler spreads with different Doppler characteristics for the different delay taps of the time-varying impulse response. An analytic technique is presented which enables the evaluation of the performance of least-squares-based coherent equalizers in such channels. The technique is based upon state-space formalisms developed in the Kalman filtering literature. Comparisons between predicted and simulated equalization performance for realistic channel models are presented.

11:15

4aUW12. Guidelines for underwater acoustic data telemetry system design. Geoffrey Edelson, Francine Menas, and Michael Richman (Advanced Systems, Sanders, A Lockheed Martin Co., P.O. Box 868, Nashua, NH 03061-0868)

The use of acoustic communications in any number of projected applications requires trade-offs within the realm of technologies currently available and those projected for the future. Analyses must be performed to predict the performance, strengths, and limitations of competing approaches. Because underwater acoustic data telemetry can be viewed as a union of sonar and communications, it suffers from the same environmental degradation as sonar and communications systems. Efforts undertaken to date have shown that system performance is dictated by received signal-to-noise ratio and multipath complexity, spread, and rate-of-change. Comparisons between different signaling technologies can be made using a variety of measures of effectiveness. These metrics provide a detailed structure for the evaluation of the performance envelope as a function of environmental and platform conditions. We have compiled published results on a large selection of modems in terms of range-rate product, bandwidth efficiency, range, frequency, and power. Based on these compiled results and on simple physical models, design rules-of-thumb have been derived to ensure that proposed system designs have communication requirements that are reasonable and robust. While these rules may seem somewhat conservative, the sample points reported in the literature may not have been achieved reliably or more than once.

Meeting of the Standards Committee Plenary Group

ORGANIZATION OF STANDARDS COMMITTEE PLENARY GROUP MEETING

• **S1 ACOUSTICS—**U.S. Technical Advisory Group (TAG) for IEC/TC 29 Electroacoustics and ISO/TC 43 Acoustics

• **S2 MECHANICAL VIBRATION AND SHOCK—**U.S. Technical Advisory Group (TAG) for ISO/TC108 Mechanical Vibration and Shock

• **S3 BIOACOUSTICS—**U.S. Technical Advisory Group (TAG) for ISO/TC 43 Acoustics, IEC/TC 29 Electroacoustics, and ISO/TC 108/SC4 Human Exposure to Mechanical Vibration and Shock

• **S12 NOISE—**U.S. Technical Advisory Group (TAG) for ISO/TC 43/SC1 Noise

The meeting of the Standards Committee Plenary Group will precede the meetings of the Accredited Standards Committees S1, S3 and S12, to take place in the following sequence on the same day:

S12	9:45 a.m. to 12 noon	California Salon 1
S3	2:00 p.m. to 3:30 p.m.	Yawl Room
S1	3:45 p.m. to 5:00 p.m.	Yawl Room

Discussion at the Standards Committee Plenary Group meeting will consist of national items relevant to all S Committees, plus a review of the international standardization (U.S. TAG) activities including reports on recent meetings and planning for forthcoming meetings.

Members of S2 on Mechanical Vibration and Shock (and U.S. TAG for ISO/TC 108 and five of its Subcommittees, SC1, SC2, SC3, SC5, and SC6) are also encouraged to attend the Standards Committee Plenary Group meeting even though the S2 meeting will take place on day earlier, on Wednesday, 6 December 2000, at 9:00 a.m.

The U.S. Technical Advisory Group (TAG) Chairs for the various international Technical Committees and Subcommittees under ISO and IEC, which are parallel to S1, S2, S3 and S12 are as follows:

U.S. TAG Chair/Vice Chair	TC or SC	U.S. TAG
ISO		
P. D. Schomer, Chair H. E. von Gierke, Vice Chair	**ISO/TC 43** Acoustics	S1 and S3
P. D. Schomer, Chair H. E. von Gierke, Vice Chair	**ISO/TC 43/SC1** Noise	S12
D. D. Reynolds, Chair	**ISO/TC 108/SC4** Human Exposure to Mechanical Vibration and Shock	S3
D. J. Evans, Chair	**ISO/TC 108** Mechanical Vibration and Shock	S2
R. H. Mehta, Chair K. Won, Vice Chair	**ISO/TC 108/SC1** Balancing, including Balancing Machines	S2
A. F. Kilcullen, Chair	**ISO/TC 108/SC2** Measurement and Evaluation of Mechanical Vibration and Shock as Applied to Machines, Vehicles and Structures	S2
D. J. Evans, Chair	**ISO/TC 108/SC3** Use and Calibration of Vibration and Shock Measuring Instruments	S2
R. L. Eshleman, Chair R. F. Taddeo, Vice Chair	**ISO/TC 108/SC5** Condition Monitoring and Diagnostics of Machines	S2
G. Booth, Chair	**ISO/TC 108/SC6** Vibration and Shock Generating Systems	S2
IEC		
V. Nedzelnitsky, U.S. TA	**IEC/TC 29** Electroacoustics	S1 and S3

Meeting of Accredited Standards Committee (ASC) S12 on Noise

P. D. Schomer, Chair S12, and Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 43/SC1, Noise
2117 Robert Drive, Champaign, Illinois 61821

R. D. Hellweg, Vice Chair, S12
Compaq Computer Corp., Acoustics Lab, Mechanical Engineering Group, MRO1-3/D3, 200 Forest Street, Marlborough, Massachusetts 01752

H. E. von Gierke, Vice Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 43/SC1, Noise
1325 Meadow Lane, Yellow Springs, Ohio 45387

Accredited Standards Committee S12 on Noise. Working group chairs will report on their progress for the production of noise standards.

Scope of S12: Standards, specifications and terminology in the field of acoustical noise pertaining to methods of measurement, evaluation and control; including biological safety, tolerance and comfort and physical acoustics as related to environmental and occupational noise.

Session 4pAAa

Architectural Acoustics: The Technical Committee on Architectural Acoustics Vern O. Knudsen Distinguished Lecture

Angelo J. Campanella, Chair
Campanella Associates, 3201 Ridgewood Drive, Hilliard, Ohio 43026-2453

Chair's Introduction—2:30

2:35

4pAAa1. Electronic architecture: Historical perspectives, current thinking, and future implications for architectural acoustic design. Paul Scarbrough (134 Washington St., D205, Norwalk, CT 06854, paulscrbro@aol.com)

Few subjects in acoustics generate more controversy than electronic architecture, that is, the use of electronic tools to alter the acoustics of a performance space, church, rehearsal hall, or other facility. And yet, few advances in acoustics offer the promise of electronic architecture. This paper will trace the development of electronic architecture from the first tentative and highly creative experiments in the 1930s to the mature technologies that have evolved in recent years. What precipitated the development of electronic architecture? Who were the pioneers? How did the technology develop? And where do we stand today? Paralleling these technological advances has been an on-going philosophical debate. Is there ever a place for electronics in the concert hall or opera house? When and where should these systems be used? What obligations do acoustical design professionals have to performers? To the concert-going public? To the art of acoustical design? Finally, what does the continuing evolution of these techniques hold for the future of architectural acoustic design?

Session 4pAAb

Architectural Acoustics: Building Acoustics Test Standards

Kenneth P. Roy, Chair
Innovation Center, Armstrong World Industries, P.O. Box 3511, Lancaster, Pennsylvania 17604

Chair's Introduction—3:45

Invited Papers

3:50

4pAAb1. Low-frequency impact sound rating of floor systems. Alf Warnock (Inst. for Res. in Construction, Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

An extensive series of measurements of airborne and impact sound insulation on joist floors of various types was recently completed at NRC. About 190 floors with different joist types, subfloors, ceiling types, ceiling support systems, and type and thickness of sound absorber were constructed and sound insulation measured to frequencies lower than required by standards. The joist types comprised solid wood, wood trusses, I-joists, and steel joists. Glass, rock, and cellulose fiber sound-absorbing material were used. Three concrete slabs were included in the series and a few of the joist floors had concrete toppings installed. Impact devices used on the floors included the standard ISO/ASTM E492 tapping machine, a walker, a JIS 1419 tire impactor, and two experimental rubber balls. The impact measurements provide an extensive, consistent set for examining alternative impact tests and possible modifications to standard tests using the ISO tapping machine. Correlations among the sound levels from the different impactors show that levels from the ISO tapping machine correlate well with those from a walker at low frequencies, although the spectra are quite different in shape. An overview of the data and the prospects for a new test method and rating system will be presented.

4:10

4pAAb2. Impact sound rating of floor toppings. Alf Warnock (Inst. for Res. in Construction, Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

Resilient materials either alone or under a slab of material are commonly placed on top of a hard floor surface to reduce impact sound transmission to rooms below. It is often impossible to select among such floor toppings when they have been tested on different base floor assemblies and only composite impact sound insulation ratings are available. ISO 140-8 measures the improvement due to a floor topping when placed on a concrete slab. The improvement may then be used to estimate the impact sound insulation of floors incorporating concrete slabs. In the project to be described, reductions in impact sound level were measured for a variety of toppings on several different floor types. The measurements confirmed that the ISO procedure works well and that relatively small specimen areas can be used without serious error for evaluation of floor toppings. The measurements showed that these improvements may not be applied to joist floors with lightweight subfloors such as plywood, but are useful for joist floors with thin concrete toppings. ASTM is adopting this test protocol for concrete slabs and is working on a modification that will provide data useful for toppings on joist floors with only light wood subfloors.

Contributed Papers

4:30

4pAAb3. Evaluating a method for field measurement of enclosure noise reduction. Felix Z. Sachs and David C. Byrne (U.S. Army Ctr. for Health Promotion and Preventive Medicine, Aberdeen Proving Ground, MD 21010-5403)

Measurement of noise reduction for sound isolating enclosures *in situ* is often complicated by limited access to all enclosure surfaces and indeterminate reverberation conditions. Both laboratory and field noise reduction measurements were made on an audiometric enclosure. The laboratory measurements followed ASTM method E596 using a reverberation chamber. The field method consisted of placing a ''calibrated'' loudspeaker at 1 m from the enclosure surface, applying a broadband signal and measuring 1 mm from the enclosure surface as well as inside the enclosure. The speaker calibration consisted of mapping the sound levels at 1 m in front of the speaker cone in an anechoic chamber. The field noise reduction is estimated using the inside and outside sound measurements. Since the loudspeaker sound field is localized, a surface area adjustment for the field noise reduction is required. The adjustment factor was derived from the E596 noise reductions. The factor is shown to be related to the loudspeaker calibration. [The opinions or assertions contained herein are the views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.]

4:45

4pAAb4. Flanking sound transmission in wood-framed construction. Trevor Nightingale, Robin Halliwell, and David Quirt (Inst. for Res. in Construction, Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

This paper presents results from the first phase of a project to characterize flanking sound transmission at wall/floor junctions in wood-framed constructions suitable for multifamily residential buildings. This phase of the study focused on the junction between a wood joist floor and a separating wall with two rows of 38×89-mm studs. The objective of the study was to identify details that simultaneously provide good noise control and fire resistance. This paper presents the experimental approach and key results from the acoustical measurements. In addition to comparing how five common fire control solutions for the floor/wall junction affect the airborne and impact sound transmission, several generic retrofit modifications to reduce the structural noise transmission were tested. Parameters studied included the floor-joist orientation (parallel or perpendicular to the

separating party wall), and the type of subfloor (plywood or oriented strand board). The study has shown that suitable details can provide both adequate fire resistance and noise control. The technical approach provides a framework for extending beyond established laboratory measurement procedures, to predict performance in actual buildings.

5:00

4pAAb5. Polish standardization of acoustic parameters of buildings in connection with ISO standards. Iwonna Zuchowicz-Wodnikowska and Anna Izewska (Bldg. Res. Inst., Ksawerow Str. 21, 02-656 Warsaw, Poland)

One of the foundations of Polish standardization in building acoustics for many years has been to establish the Polish Norms, as EN ISO Standards translated into the Polish language. In recent time, following establishment of ISO 717, it became necessary to amend the Polish Standard (PN-B-02151-3:1999) in which the required acoustic insulation were defined. The new quantities of air-borne noise insulation are adapted to two model types of noise spectrum: pink noise characteristic, e.g., for residential noise, and road traffic noise with a considerable content of acoustic pressure levels in low- and medium-frequency ranges. The standard provides the following basic rules to determine the apparent indices of acoustic insulation (depending on the model spectrum): $R'A1 = R'W + C$ or $R'A2 = R'W + Ctr$. The standard also includes an informative supplement describing the simplified computing method of the flanking transmission. The key advantage of this standard lies in its obligatory character, i.e., it must be adhered to in both the design as well as execution of buildings, and it ensures the basic standard of acoustic performance. To cater to the needs of the designers, the Acoustic Division of ITB prepared a classification of building products comprising the indicated standard for measurement or computation and the class symbol and the range of the single-number quantity.

THURSDAY AFTERNOON, 7 DECEMBER 2000 SCHOONER/SLOOP ROOMS, 2:00 TO 5:10 P.M.

Session 4pAB

Animal Bioacoustics: William Cummings Session on the Acoustics of Whales and Dolphins I

D. Vance Holliday, Chair
BAE Systems, 4669 Murphy Canyon Road, Suite 102, San Diego, California 92123

Chair's Introduction—2:00

Invited Papers

2:10

4pAB1. First known study, identification, and characteristics of fin and blue-whale vocalizations in the Northeast Pacific with historical notes and anecdotes. William C. Cummings (Oceanographic Consultants, 5948 Eton Ct., San Diego, CA 92103, oshendbs@aol.com) and Paul O. Thompson (San Diego, CA 92103)

In 1964 we were tasked by the Navy to identify and describe low-frequency, high-level noise mainly in the 1/3rd-octave frequency band containing 20 Hz. Due to AGC and other presently antiquated processing then used at Northeast Pacific sound surveillance stations (SOSUS), acoustical monitoring was often confounded for significant annual periods heightening concern during that era of Cold War. Using data from continuous recordings for Gordon Wenz's ambient noise studies, we characterized the causative individual sources mainly as trains of approximately 1-sec pulsed pairs and as long series of about 20-sec paired moans. We determined seasonality, noise acceleration, geographic variation, representative source levels, and identity of the sources which turned out to be whales. This paper describes methods and findings of several years' study to identify and describe this ''noise'' employing numerous expeditions by sea and aircraft, now a more than 30-year-old precedent effort for which the authors (or their colleagues whose contributions shall be specified) have never been fully recognized by the military or by contemporary scientists. Intermingled we include trivia, such as the fate of the invaluable Wenz recordings, and interesting accounts of simultaneous Northwest Atlantic investigations by Schevill, Watkins, and colleagues including the present first author.

2:30

4pAB2. Characteristics of chorusing sounds of humpback whales wintering in waters off western Maui. Whitlow W. L. Au, Marc O. Lammers, Paul E. Nachtigall (Hawaii Inst. of Marine Biol., Univ. of Hawaii, P.O. Box 1106, Kailua, HI 96734), Joseph Mobely (Univ. of Hawaii–West O'ahu, Honolulu, HI), and William C. Burgess (Greenride Sci., Inc., Santa Barbara, CA)

The dominant biological sound in the waters of west Maui during the humpback whale winter season in Hawaii is the chorusing sounds of singing humpback whales. A portable data logger was used to record chorusing sounds during most of the 1998 season. Sounds were digitized for 4 min every $\frac{1}{2}$ h and the data were stored on a hard disk. The results showed a peak in the sound pressure level between mid-February and mid-March coinciding with the peak in the number of whales sighted by aerial surveys. Significantly higher levels of chorusing sounds were also recorded at night. Sound levels began to increase during sunset, raising to a peak near midnight and remained relatively high until sunrise, after which the levels progressively decreased to a minimum between 11:00 a.m. and 3:00 p.m. That more humpback whales appear to sing at night may reflect a switch to sexual advertisement as the primary male mating strategy at this time; it may also indicate that daylight and vision play key roles in the formation of competitive groups. It is suggested that the relative number of humpback whales may be estimated by monitoring changes in sound pressure levels.

2:50

4pAB3. Three-dimensional localization of blue whales using matched-field processing techniques: Unusual behavior, unanswered questions. A. M. Thode (MIT, 77 Massachusetts Ave., Cambridge, MA 02139), G. L D'Spain, and W. A. Kuperman (Scripps Inst. of Oceanogr., San Diego, CA 92093)

Recently, matched-field processing (MFP) and global inversion methods were used to track the three-dimensional trajectories of six blue whales, during a 1996 experiment off San Miguel Island in California, conducted off the research platform FLIP, using a 48-element, tilted vertical hydrophone array. Once the animals' positions and the propagation environment were determined, source levels were obtained for individual animals, and potential propagation effects were removed from their sounds. Out of the six whales studied, one showed unusual behavior. Initially it swam directly towards FLIP, while producing a diverse set of calls that included three strong FM sweeps. It then fell silent, altered its course to avoid FLIP, and began calling again as it swam away. Its closest point of approach to FLIP coincided with the arrival of a powerful tidal bore. The maximum source levels estimated from this animal were 10 dB lower than those from the other animals, and the FM sweeps showed signs of a possible resonance, which remained even when possible propagation effects were removed. The ability to precisely position blue whales has raised interesting questions about baleen whale sound production, the purpose behind these sounds, and the potential association between feeding animals and internal waves. [Work sponsored by ONR.]

3:10

4pAB4. Seasonal contribution of mysticete vocalization to ambient noise in southern California waters. David A. Helweg (Code D351, SSC San Diego, 49620 Beluga Rd., San Diego, CA 92152, helweg@spawar.navy.mil)

Many species of baleen whales reside or migrate annually through California waters. During these times some species of vocalizing whales can be detected, identified, and tracked with SOSUS arrays. Historical research by W. Cummings documented strong seasonal changes in narrow-band ambient sound associated with then unidentified sounds. Contemporary research shows these sounds to be blue and fin whale vocalizations. The current study compares historical and contemporary narrow-band ambient measurements with respect to mysticete residency and movement through southern California waters. Seasonal peaks in ambient noise agree with earlier reports, but increased anthropogenic sound reduced the magnitude of the effect. Potential sources of changes in ambient noise and their effects on the whales are discussed. [Research supported by the Strategic Environmental Research and Development Program, Project CS-1082.]

3:30

4pAB5. Thirty years of Australian marine bioacoustics and the Cummings connection. Douglas H. Cato (Defence Sci. and Technol. Organisation, P.O. Box 44, Pyrmont, NSW 2009, Australia), Robert D. McCauley (Curtin Univ. of Technol., Bentley, Western Australia), Michael J. Noad, and Tracey Rogers (Univ. of Sydney, NSW 2006, Australia)

Marine bioacoustics research in Australian waters has developed for more than 30 years, and has revealed a wide diversity of sounds and acoustic environments. The number of researchers and the complexity of the experiments has increased over the years. In the early days we had little knowledge or experience, and so relied on advice and help from American researchers. Of those, Dr. William C. Cummings remains one of our greatest supporters. This paper reviews some of the main features of the Australian work and the Cummings influence. Some of the earliest work, on fish sounds, echoes his work of the 1960s. This has since developed substantially, and we are planning to apply methods developed by Cummings of combined video and acoustic observation to source identification. Humpback whale sounds have been studied continually for more than 20 years and have revealed the complex songs of similar overall structure to those of the northern hemisphere, but with some surprises. Recordings have been made of southern right whales and Antarctic seals, and this year we recorded a large set of diverse sounds from blue whales. This data set and visual / acoustic observations of humpback whales are being used in censusing.

3:50

4pAB6. Two independent sonar signal generators in the bottlenose dolphin: Physiologic evidence and implications. Ted W. Cranford (Dept. of Biol., San Diego State Univ., San Diego, CA 92166), Wesley R. Elsberry, Diane J. Blackwood (Texas A&M Univ., College Station, TX), Jennifer A. Carr, Tricia Kamolnick, Mark Todd (Science Applications Intl. Corp., San Diego, CA), William G. Van Bonn, Donald A. Carder, Sam H. Ridgway (U.S. Navy Marine Mammal Prog., San Diego, CA), Devon M. Bozlinski, and Emily C. Decker (Univ. of San Diego, San Diego, CA)

Some indirect evidence exists for multiple sonar signal generators in odontocetes. Direct evidence was collected from two bottlenose dolphins by simultaneously measuring and digitizing internal nasal cavity pressure, nasal tissue motion, and acoustic pressure. Small catheters measured pneumatic pressure changes at the same depth within each bony nasal passage. A high-speed video endoscopy system revealed tissue motion at both pairs of phonic lips while two hydrophones measured acoustic pressure during biosonar target discrimination. The records clearly demonstrate that acoustic pulses can be generated at the phonic lips on the left and right sides, independently or simultaneously. We have only seen the left phonic lips generate whistles. Air pressure in the bony nasal passages rises and falls together, even if the activity patterns at the two pairs of phonic lips are different. Increasing pulse repetition rate or sound pressure level, and whistle production, are all normally accompanied by increasing nasal air pressure. Whistle production apparently requires greater nasal air pressure than does pulse generation. Discovering multiple sonar sources in a dolphin's head should cause us to reinterpret previous findings and re-evaluate physiologic limits on pulse repetition rate, signal bandwidth, frequency composition, and projection beam geometry. [Work sponsored by Office of Naval Research.]

4:10

4pAB7. High-frequency auditory filter shapes in an Atlantic bottlenose dolphin. David W. Lemonds, Whitlow W. L. Au, Paul E. Nachtigall, Herbert L. Roitblat (Marine Mammal Res. Prog., Hawaii Inst. of Marine Biol., Univ. of Hawaii, P.O. Box 1106, Kailua, HI 96734, lemonds@hawaii.edu), and Stephanie A. Vlachos (Sea Life Park Hawaii, Waimanalo, HI 96795)

High-frequency auditory filter shapes were calculated for a 20-year-old female Atlantic bottlenose dolphin (*Tursiops truncatus*). Thresholds were determined for tones of 40, 60, 80, and 100 kHz masked by notched noise. Auditory filter shapes were determined by fitting the integral of the roex(p,r) filter shape to the functions relating masked tonal threshold to notch width. Filter shapes were found to be approximately symmetric at the moderate noise level used. Equivalent rectangular bandwidths of the auditory filters ranged from 16% of center frequency at 40 kHz to 11% of center frequency at 100 kHz. There was very little change in the bandwidths of the filters between 60 and 100 kHz, indicating that relative tuning sharpness increases as a function of frequency in this range. Efficiency of processing after the periphery was found to be maximal at 40 and 60 kHz (better than 12-dB SNR) and to decrease gradually above 60 kHz. The efficiency estimates allowed for the reanalysis of critical ratio data collected previously for the same animal; the bandwidth estimates from the present and earlier study were found to be in good agreement.

4:30

4pAB8. Foraging behavior of the Hawaiian spinner dolphin observed with active acoustics. Kelly Benoit Bird and Whitlow Au (Hawaii Inst. of Marine Biol., 46-700 Lilipuna Rd., Kaneohe, HI 96744)

The mesopelagic boundary community (MBC) in the Hawaiian Islands is an important food resource for many organisms, including spinner dolphins (*Stenella longirostris*). To understand the role of this community in the feeding ecology of spinner dolphins, spatial and temporal patterns in the density of the MBC and relative abundance of dolphins were characterized using acoustic sampling. A 200 kHz echosounder modified to sample directly into a computer was used to survey the leeward coast of Hawaii from 2100 until 0400. Acoustic data were mapped in ArcView to show the distribution of both the predator and its prey. Individual echoes were color-coded by target strength to allow the separation of spinner dolphins and mesopelagic organisms. Dolphin species identification from active acoustics was confirmed using a combination of passive acoustics and visual observations. The MBC migrated horizontally, raising questions about the accepted (Norris) model of spinner dolphin foraging in Hawaii. Observations of spinner dolphins suggest a more complex foraging pattern, compatible with patterns observed in their food resource. Spinner dolphin foraging patterns closely matched the spatial patterns of their prey both horizontally and vertically. Patches exploited by dolphins were also significantly different than other patches measured under similar conditions.

4:50

4pAB9. Blue whale calls recorded in the Gulf of Alaska. Kathleen Stafford (CIMRS/COAS Hatfield Marine Sci. Ctr., Newport, OR 97365)

Blue whales (Balaenoptera musculus) were known to range throughout the North Pacific Ocean, including the Gulf of Alaska, before whaling activities greatly reduced their numbers. Despite repeated dedicated aerial and shipboard surveys in the Gulf of Alaska and off the Aleutian Islands, very few blue whales have been seen in these waters [NMFS, Mar. Fish. Rev. **61**, 3843 (1999)]. Recent studies have shown the utility of acoustic methods for monitoring blue whale calls in the northeastern Pacific [cf. Thompson and Friedl, Cetology **45**, 1–19 (1982); Stafford *et al.*, Mar. Mamm. Sci. **15**, 1258–1268 (1999); Watkins *et al.*, Oceanography **13**, 62–67 (2000)] but none of these have surveyed the Gulf of Alaska. In order to address the question of whether or not blue whales are still found in the Gulf of Alaska, six hydrophones were deployed there in October 1999. The first data set was recovered in May 2000. A preliminary analysis of these data showed that both northeastern [cf. Rivers, Mar. Mamm. Sci. **13**, 186–195 (1997)] and northwestern [cf. Thompson and Friedl (1982)] Pacific blue whale calls were recorded on the hydrophones.

Session 4pAO

Acoustical Oceanography: Acoustic Measurements of Sediment Transport

Alex E. Hay, Chair
Department of Oceanography, Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada

Chair's Introduction—1:00

Invited Papers

1:05

4pAO1. TRIDISMA: An EU program for utilizing acoustics for measuring suspended sediment processes. Peter D. Thorne (Ctr. for Coastal and Marine Sci., Proudman Oceanographic Lab., Bidston Observatory, Bidston, Prenton CH43 7RA, UK), Jon J. Taylor (Env. Sci., Univ. of East Anglia, UK), Victor F. Humphrey, and Christopher E. Vincent (Bath Univ., UK)

In this paper the application of high-frequency acoustic remote sensing is discussed with respect to the development of an instrument capable of resolving the dynamics of the benthic boundary layer, in response to the interaction of waves and currents, using sound backscattered by the sediments entrained by the flow. Within the TRIDISMA project an acoustic instrument has been developed which combines three distinct approaches to the analysis of backscattered signals in order to extract information on flow structure and the response of the sediment bed to the overlying flow. The TRIDISMA instrument combines pulse coherent Doppler and acoustic cross-correlation techniques to the problem of extracting profiles of velocity, while the amplitude of the backscattered signal measured at three frequencies is used to investigate the response of the suspension field in terms of both size and concentration profiles. In the present paper we describe the hardware and software architecture of the prototype instrument and demonstrate the systems capabilities with results obtained from flume studies of waves, currents, and combined flows.

1:30

4pAO2. Rake, a wave measuring benthic acoustic stress sensor. Archie T. MorrisonIII (McLane Research Labs., Inc., Falmouth Technology Park, 121 Bernard E. Saint Jean Dr., East Falmouth, MA 02536, atmorrison@mclanelabs.com) and Albert J. WilliamsIII (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

The Rake is an acoustic travel time current meter designed to make spatially and temporally dense velocity profile measurements in the continental shelf wave bottom boundary layer. The high levels of bottom shear stress generated in this extremely thin, oscillatory, boundary layer play an important role in the entrainment of sediment into the water column and also enhance the turbulent dissipation of energy in overlying steady flows. The array of acoustic transducers measures the horizontal velocity vector with millimeter vertical resolution at up to 64 heights. A multiplexing transmit/receive circuit permits measurement of the vertical component of velocity with 5 to 10 cm vertical resolution using the same array. Complete velocity profiles can be acquired within a 40 ms window. Measurement error due to velocity averaging over the sample volume of a single acoustic axis has been shown to be negligible to within 1 mm of the bottom both by calculation and experiment. Laboratory calibrations and experiments have been conducted in steady and oscillatory flow. A month-long deployment of a field prototype took place in December of 1996. Data from these analyses and experiments will be presented. [The development of the Rake was supported by ONR and NSF.]

1:55

4pAO3. Acoustic measurements of boundary layer turbulence, sediment concentration and sediment flux using the Bistatic Coherent Doppler Velocity and Sediment Profiler (BCDVSP). Tim Stanton (Oceanogr. Dept., Code OC/St, Naval Postgrad. School, Monterey, CA 93943)

A pulse-to-pulse coherent acoustic doppler using a bistatic transducer configuration has been developed to study oceanic boundary layers. The instrument uses a central narrow-beam transceiving transducer, with three surrounding fan-beam response receiving transducers directed toward the narrow volume ensonified by the central transducer. Short acoustic pulses ensonify approximately 1-cm-long sections of the water column below the 2-deg-width transmitter transducer, and acoustic energy reflected from scatterers in the water column are received by all four transducers. Coherent sampling methods provide high resolution doppler frequency measurements. The four doppler estimates allow three component velocity vectors to be estimated concurrently with the backscatter amplitude in the same small volumes over a 75-cm range, at 1-cm resolution and 20-Hz sample rate. Laboratory calibration of the acoustic backscatter using *in situ* sand samples provides sediment concentration profiles over a 0.001 to 25 g/l concentration range. The BCDVSP has been used in two ONR studies of the oceanic bottom boundary layer (BBL) on the inner shelf. Examples of BBL turbulent structure and sediment concentration will be shown from the SandyDuck near shore experiment in 1997, and from two inner shelf sites during the Shoaling Waves Experiment.

2:20

4pAO4. Coherent Doppler profiler measurements of near-bed suspended sediment fluxes for different bed states. Carolyn E. Smyth, Alex E. Hay (Dalhousie Univ., Halifax, NS B3H 4J1, Canada), and Len Zedel (Memorial Univ., St. John's, NF)

Vertical profiles of vertical turbulence intensity and vertical sediment fluxes were collected by an acoustic coherent Doppler profiler at two locations: Queensland Beach (Nova Scotia), and Duck (North Carolina). Observations of the turbulence intensity over different bed states (irregular ripples, cross ripples, linear-transition ripples, and flat bed) reveal that the near-bed turbulence levels are strongly affected by bed forms. This study examines the mechanisms of distributing suspended sediments and generating near-bed turbulence for the four bed states based on the characteristics of two previously observed mechanisms: diffusion and vortex shedding. Wave-phase averages of turbulence intensity, suspended sediment concentration, and suspended sediment fluxes are compared to vortex shedding and diffusion signatures. Evidence of vortex shedding is found for the low-energy ripples, but no signatures of diffusion are observed. Two diffusion models including a bed stress model, and an eddy diffusion model are found to predict near-bed turbulence levels reasonably well, but predictions are inconsistent with the trend of the data. A vortex-shedding model [J. F. A. Sleath, J. Fluid Mech. **182**, 369–409 (1987)] predicts the vertical structure of the turbulence for rippled beds when the ripple wavelength is used as a ripple roughness.

2:35

4pAO5. Analysis of acoustic measurements of suspended sediments collected in a large-scale wave flume facility. Peter D. Thorne (Proudman Oceanographic Lab., Bidston Observatory, Bidston, UK)

Our capability to accurately predict suspended sediment concentration in the marine environment is relatively limited; generally, measuring techniques to obtain suspended concentration have been somewhat primitive. In recent years acoustics has been applied to the measurement of suspended sediments, and through a series of theoretical and experimental studies, a comprehensive description of the interaction of sound with suspensions of marine sediments has been developed. In the present work these developments have been applied to obtain detailed acoustic profiles of suspended sediment concentration under waves in a large-scale flume facility. Utilizing the acoustic concentration profiles, an assessment of suspended sediment models for predicting sediment concentration has been conducted, and the underlying intrawave mechanisms of sediment entrainment investigated.

2:50

4pAO6. Three-dimensional imaging of sand transport processes over wave ripples. Peter Traykovski, James D. Irish (Appl. Ocean Phys. and Eng. Dept., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Alex E. Hay (Dalhousie Univ., Halifax, NS B3H 4J1, Canada)

While the interactions between bedforms, suspended sediment transport and the forcing hydrodynamics are inherently three dimensional, most observations have been either point samples as a function of time, or 1-D vertical imaging systems. Recently, acoustic imaging systems have been designed to make 2-D maps of either bedform plan view topography (rotary fan beam sonars) or suspended sand concentration and bedform profiles (rotary pencil beam sonars). In the winter of 1999–2000, Woods Hole and Dalhousie investigators combined several of these 1- and 2-D systems to produce a data set on the 3-D structure of the suspended sand concentration field and its relation to the seafloor topography and forcing hydrodynamics. The data revealed that although the wave orbital scale ripples can be quite 2-D at times (i.e., long crested), the suspended sediment has considerable structure in the along-crest direction. The bedforms were observed to migrate in the onshore direction, and clouds of suspended sand were observed to be re-suspended in the lee of a ripple crest and then were advected back and forth over the crest.

3:05–3:20 Break

3:20

4pAO7. Field observations of bedforms in high sediment transport conditions. Anna M. Crawford and Alex E. Hay (Dept. of Oceanogr., Dalhousie Univ., Halifax, NS B3H 4J1, Canada)

In the first reported use of MHz-frequency acoustics to study small-scale bedforms in the near shore zone, Dingler and Inman were able to observe the evolution of transition ripples under energetic groupy waves near the flat bed threshold. They found that these low-relief, long-crested bedforms appeared to be obliterated under the largest waves in a group, and subsequently reformed in one or two wave periods [J. Dingler and D. Inman, Proc. 15th Coastal Conf., Vol. II, 2109–2126 (1976)]. More recently, measurements of linear transition ripples were made during an autumn storm event using scanning wide beam and fixed narrow beam sonars, as well as a high resolution laser-illuminated video system. These observations also indicate the occasional apparent disappearance of these ripples, and also their reformation at the same locations. The question arises as to whether the disappearance is apparent (the result of increased near bed sediment concentration and the definition of bed elevation based on reflection amplitude) or real (planing off of the ripple crests by sheet flow).

3:35

4pAO8. A three-dimensional coherent Doppler profiler. Len Zedel (Phys. and Physical Oceanogr. Dept., Memorial Univ., St. John's, NF A1B 3X7, Canada, zedel@physics.mun.ca) and Alex Hay (Dalhousie Univ., Halifax, NS B3H 4J1, Canada, hay@phys.ocean.dal.ca)

A 1.7-MHz coherent Doppler sonar system has been designed for near-shore environments that provides three-dimensional velocity profiles over an O (0.5-m) depth range. The use of a bistatic geometry allows three independent components of velocity to be measured simultaneously in time and coincident in space. The system was calibrated in a tow-tank facility for velocities up to 2 m/s. For multiple pulse-pair ensemble-averaged velocity estimates generated at a rate of 30 profiles/s in 0.7-cm-depth bins, vertical velocity accuracies of the order 1 cm/s are achieved. For the horizontal components (based on bistatic measurements), standard deviation in velocity estimates of about 5% is realized with absolute accuracies of about 2%. In the configuration as tested, flow disturbance around the instrument introduces a bias of about 5%: this disturbance is modeled using potential flow past a cylinder. The disturbance errors could be corrected to the free-stream velocities using the present data set. Such a correction would not, however, reflect the actual water speed at the point of measurement. [This research was funded by the Office of Naval Research Coastal Sciences Program, and the Natural Sciences and Engineering Research Council of Canada.]

3:50

4pAO9. Effects of the top transition layer in marine sediments on seabed roughness and volume scattering. Anatoliy N. Ivakin (CNRS/LMA/ASM2, 31 Chemin Joseph-Aiguier, 13402 Marseille Cedex 20, France; also affiliated with AOPE/WHOI, Woods Hole, MA 02543, anivakin@hotmail.com)

Most models of high-frequency seabed scattering normally assume that the sediment is a statistically homogeneous half space and there is a discontinuity of acoustical parameters at the water–sediment interface. It is known, however, that in many cases properties of the very top layer of real marine sediments actually represent almost continuous transition within a few centimeters from those in water to those essentially constant in sediments at greater depths. At the Key West site, e.g., a definite transition layer in silty sand was found to be of 6-cm thickness. The discontinuity at the water–sediment interface, if it exists, is actually much smaller than assumed in conventional models of high-frequency scattering, where roughness scatter component is usually predicted to be strong

and dominating for most types of sediments except very fine ones, such as mud and fine silt. In this paper, effects of the transition layer on both roughness and volume-scattering components are considered. These effects are shown to reduce substantially the role of roughness scattering at high frequencies. In particular, the roughness contribution is shown to be strongly dependent on real discontinuity of acoustical parameters at the water–sediment interface, which can be significantly reduced even for some sands due to existence of the transition layer. The possibly dominating role of volume scattering caused by continuous and discrete heterogeneity of the sediment is discussed. The necessity of more detailed study of physical properties of the transition sediment layer is emphasized. Related issues of interpretation of recent results on seabed scattering obtained in the Key West and SAX99 experiments are discussed as well. [Work supported by ONR-US and CNRS-France.]

THURSDAY AFTERNOON, 7 DECEMBER 2000 CALIFORNIA SALON 2, 1:25 TO 4:30 P.M.

Session 4pMUa

Musical Acoustics: Historical Brasses and Related Instruments

R. Dean Ayers, Chair
Department of Physics and Astronomy, California State University, Long Beach, California 90840-3901

Chair's Introduction—1:25

Invited Papers

1:30

4pMUa1. An overview of the brass collection in the Kenneth G. Fiske Museum. Albert Rice (Kenneth G. Fiske Museum of The Claremont College, Claremont, CA 91711, alrice@cuc.claremont.edu)

The Kenneth G. Fiske Museum of Musical Instruments at The Claremont Colleges is one of only seven comprehensive musical instrument museums in the United States. Its collection of brasses and cup-mouthpiece instruments made of wood consists of over 400 instruments dating from the 18th through the 20th centuries. It includes cornets, trumpets, bugles, trombones, French horns, mellophones, euphoniums, tenor horns, ophicleides, keyed bugles, over-the-shoulder horns, serpents, cornettos, Russian bassoons, bass horns, tubas, instruments from Africa, Tibet, China, and other countries. The collection is especially rich in its representation of important 19th century American makers and of examples of valve designs such as Stoelzel, Berlin Pumpen, Vienna, Allen, and various designs of the Perinet valve. Valves were initially introduced by German makers in 1818, and were gradually accepted by players during the 1830s and 1840s. The Perinet valve of 1838 emerged as the standard design only by the 1870s and 1880s. With the use of slides of several instruments the presentation will illustrate four broad aspects of the collection: (1) the development of the low brasses; (2) types of valves; (3) brasses in Asia and other countries; and (4) unusual designs.

2:00

4pMUa2. The evolution of lip-reed instruments and their manufacture. Robert W. Pyle, Jr. (11 Holworthy Pl., Cambridge, MA 02138)

Lip-reed instruments began in prehistory as animal horns and tusks and have evolved into the brass instruments of modern orchestras and bands. After a brief look at instruments of biblical times and earlier, this paper will concentrate on developments from the 15th century to the present. By the mid-18th century, the typical trumpet bell contour had changed from an ''animal horn'' shape to something very close to that of the present-day trumpet and players had achieved remarkable technical proficiency, as shown in the compositions of Bach, Handel, and others. On the other hand, manufacturing techniques had changed hardly at all and the instruments were still mechanically simple. The introduction of steam power in the late 18th century profoundly changed metal-working techniques and, with the invention of the valve, led to an explosion in the variety and complexity of brass instruments. Of the numerous dead branches and twigs on the brass family tree of the 19th century, some were killed by bad acoustics, some by economics, and some by poor marketing. In our own day, computer-controlled machining not only has changed traditional manufacturing economics, but also has allowed the production of new components that would formerly have been prohibitively expensive.

2:30

4pMUa3. An acoustical comparison of the serpent and the ophicleide. D. Murray Campbell (Dept. of Phys. and Astron., Univ. of Edinburgh, Edinburgh EH9 3JZ, UK)

The serpent is a lip-excited wind instrument with an air column length of around 2 m and an approximately conical bore. In the form in which it emerged in France in the sixteenth century, it is a wooden tube with six finger holes and a chromatic compass of three octaves. The serpent suffers from some problems related to the fact that all the holes are covered directly by the fingers. The holes are grouped in two clusters, and the hole diameter is too small to provide adequate venting. The ophicleide, invented in the early nineteenth century, has a similar bore profile to the serpent. Eleven or more side holes in the thin-walled metal tube are covered by lever-operated padded keys; these allow most notes to be obtained without cross fingering, and the size and placing of the holes are not limited by the anatomy of the human hand. Several acoustical investigation techniques, including input impedance analysis, pulse reflectometry and excitation by an artificial mouth, have been used to investigate the extent to which the ophicleide is a successful solution to the acoustical problems posed by the serpent.

3:00

4pMUa4. A survey of the historical and acoustical development of the horn. Brian Holmes (Dept. of Phys., San Jose State Univ., San Jose, CA 95192-0106)

In the 17th century, the hunting horn entered the orchestra. With its shallow mouthpiece and rapidly flared bell, the baroque horn had a bright sound like that of the trumpet. Indeed, the same musicians often played either instrument. In modern authentic performances, players use side-vents or hand-stopping to improve intonation, though no evidence suggests that 18th century players used these techniques. When the hand entered the bell, the bell was widened and the mouthpiece deepened, giving the instrument its modern character. Despite the invention of the valve, acoustics requires that the hand remain in the bell. The modern double horn, with its wider bell and rotary valves, was introduced by Kruspe in 1902. Older designs (such as the Vienna horn, with its narrower bell, terminal crook, and curious Vienna valves) persist in some areas. The drive to compete with trumpets for volume and recordings for accuracy has led some modern players to adapt higher pitched descant horns.

3:30

4pMUa5. Design characteristics of nineteenth century American-made brass instruments. Richard P. Birkemeier (Dept. of Music, California State Univ., Long Beach, 1250 Bellflower Blvd., Long Beach, CA 90840, birkemei@csulb.edu)

The brass band was arguably the most popular form of entertainment for most American citizens in the nineteenth century. This ensemble flourished in nearly every town and eventually led directly to the development of jazz, the basis for all American popular music in the twentieth century. American brass instrument design underwent significant changes in a relatively short period of time during the second half of the nineteenth century. These design changes resulted in noticeable variations in instrumental timbre, changes that clearly had an effect on the music written for wind bands during this period. This presentation will feature two complete sets of nineteenth century brass band instruments, one from the Civil War era and one from the end of the century. They will be analyzed from the following perspectives: design specifications of bore, bell shape, valve configuration and mouthpiece design; weight and thickness of metal; and instrument timbre as demonstrated in live performance. It is hoped that the comparison of design characteristics followed by the actual sounds of the instruments will offer insights into this most important of relationships for the musician.

Contributed Papers

4:00

4pMUa6. Pitch drift as a result of just intonation. Patrick K. Mullen (Art Dept., CSU Long Beach, Long Beach, CA 90840)

The most easily recognized difficulty with just tuning is the large and confusing array of pitches that must be made available on a keyboard instrument in order to allow modulation to several keys. A given note name may have more than one pitch associated with it, sometimes several, depending on the context in which it is used. Using computer-generated instruments that provide a limitless supply of unique pitches, just intervals can be used exclusively and explored thoroughly. When applied to Giovani Gabrieli's Sonata Pian' E Forte, the strict use of just tuning forces the pitch center to drift lower as the music progresses. When pitches are held across certain chord changes, lower versions of previous pitches must be used in order to maintain just tuning. Though most nonkeyboard acoustic instruments, the brasses in this case, can vary the tuning of each pitch somewhat, most cannot match the degree of drift that occurs in the Gabrieli (''A'' begins the piece at 440 Hz and ends it at $\sim$398 Hz!). For players of these instruments a compromise between physical tuning limitations and the ideal of just intonation must be made. Players use just intervals wherever possible, until the pitch drift forces a decision as to where the ''rub'' will be. The pitch-drift phenomenon will be demonstrated.

4:15

4pMUa7. Musical instruments of antiquity as illustrated in *The Adventures of Asterix the Gaul*. Daniel A. Russell (Sci. & Math Dept., Kettering Univ., 1700 W. Third Ave., Flint, MI 48504)

The Adventures of Asterix the Gaul, a series of books written by Goscinny and illustrated by Uderzo, between 1960 and 1999, have received much acclaim for the attention to detail in their drawings of ancient architecture and civilization. Equally interesting to a musical acoustician are the illustrations of musical instruments (including carnyx, buccina, lur, bagpipe, harp, lyre, pipes, drums, and singing) used by ancient Romans, Greeks, Egyptians, and Gauls. This talk will compare Uderzo's illustrations to photographs of period instruments and comment on their acoustic qualities, performance techniques, and the roles they played in their respective societies, both in real history and as experienced by Asterix and his friends.

Session 4pMUb

Musical Acoustics: Concert Performance by the Americus Brass Band

R. Dean Ayers, Chair
Department of Physics and Astronomy, California State University, Long Beach, California 90840-3901

Chair's Introduction—5:00

A performance technical session featuring the Americus Brass Band will be held on Thursday from 5:00 to 6:00 p.m. in the California Ballroom. The band re-creates the historical town band of Americus, Georgia, which became the regimental band of the 4th Georgia Volunteers in the Confederate Army. The Americus Brass Band has toured nationally for eight years, appearing at major Civil War battle reenactments, balls, and other related ''living-history'' events. They have produced three critically acclaimed CDs and recorded sound-track music for and/or appeared on-screen in a dozen major motion pictures and television shows. Playing their antique instruments and dressed in period costume, the Americus Brass Band will offer a rousing concert of music from the Civil War to the Buffalo Bill Wild West Show, demonstrating how brass instruments and brass music evolved in the United States during the last half of the nineteenth century. Visit them at: http://www.csulb.edu/~birkemei/abbweb/abb.htm

Session 4pNS

Noise: Hearing Protection Standards

Elliott H. Berger, Chair
E-A-R, 7911 Zionsville Road, Indianapolis, Indiana 46268-1657

Chair's Introduction—1:30

Invited Papers

1:35

4pNS1. New ISO standards for hearing protectors. Torben Poulsen (Dept. of Acoust. Technol., Bldg. 352, DTU, DK-2800 Lyngby, Denmark, tp@dat.dtu.dk)

Working Group 17 under ISO/TC43/SC1 deals with measurement of hearing protector attenuation. WG17 is presently involved in revision of the present 4869-1 standard (especially sound field requirements and reproducibility estimations) and formulation of a new standard where inexperienced hearing protector users are mandatory as test subjects. The latter procedure has recently been introduced in the USA, Australia, and New Zealand. Both the present and the new standard use the Real Ear At Threshold (REAT) method. Results from investigations of the sound field in different test environments will be presented and the status of coming standards for impulse noise and for devices with active noise control will be mentioned.

1:55

4pNS2. Ongoing experience with laboratory-based subject-fit REAT methods for measuring hearing protector attenuation. Elliott H. Berger and Ronald W. Kieper (E-A-R, 7911 Zionsville Rd., Indianapolis, IN 46268-1657)

As data have accumulated indicating the difficulty in basing estimates of field-delivered (i.e., real-world) protection for groups of users upon classical laboratory-derived optimum attenuation data, more attention has been directed towards standards intended to provide better simulations of the real world. The development of ANSI S12.6-1997 *Methods for Measuring the Real-Ear Attenuation of Hearing Protectors*, is one approach towards addressing this problem. The standard includes both experimenter-supervised fit and subject-fit methods. The latter method, designated as Method B, requires the use of audiometrically experienced subjects who are naive in the use of hearing protection. This method is intended to approximate the upper limits to the attenuation that can be expected for *groups* of occupational users. It yields mean attenuation values, more so for earplugs than earmuffs, that are substantially lower

and standard deviations values that are higher than previously found using ANSI standards. This paper will present the latest Method-B test data, review ongoing studies designed to further evaluate the Method-B protocol, and will comment on efforts to more widely disseminate Method-B data and the associated noise reduction rating (Subject FIT) [NRR(SF)] which can be derived from such values.

2:15

4pNS3. Measurement and classification of hearing protectors in New Zealand and Australia. George Bellhouse (Bel Acoustics Consulting, P.O. Box 22 073, Khandallah, Wellington, New Zealand, george.bellhouse@xtra.co.nz)

An overview of the method of the measurement of the attenuation of hearing protectors using the subject-fit protocol according to the Joint Australian/New Zealand Standard AS/NZS 1270-1999: ''Acoustics—Hearing Protectors'' is presented including a comparison with ANSI S12.6-1997: ''Methods of Measuring the Real-Ear Attenuation of Hearing Protectors''; Method ''B'': Subject Fit. In this Australian/New Zealand Joint Standard, hearing protectors are rated in terms of a ''class'' rather than given an attenuation value. The system of classifying hearing protectors is detailed including the actual method of the determination of the ''class'' of a hearing protector. The use of the Australia/New Zealand ''class'' system in the workplace is illustrated by way of an example. This describes the use of LAeq,T, the determination of LAeq,8h and how these parameters relate to the ''class'' of a hearing protector. The benefits and drawbacks of the system used are compared with other rating systems.

2:35

4pNS4. Now what do we do with these good numbers? John R. Franks and William J. Murphy (Hearing Loss Prevention Section, NIOSH, M.S. C-27, 4676 Columbia Pkwy., Cincinnati, OH 45226-1998)

Present ANSI and ISO standards suggest calculating the mean real-ear-attenuation at threshold (REAT) and the standard deviations as well. These are used for calculating assume protection values (APVs) that are incorporated in calculations of the NRR in the United States and the SNR and HML in Europe and elsewhere. One of the discoveries from intense analysis of both experiment-fit and subject-fit REAT values is that they are not always unimodally and normally distributed around a mean or median, thus weakening the statistical relevance of the mean and standard deviation of the REATs. Methods for determining APVs that take into account the distribution of REATs and yet allow the computations to be otherwise unaffected will be presented. Additionally, suggestions for ratings other than the NRR, SNR, and HML will be made.

2:55

4pNS5. Measuring the noise attenuation of shotblasting helmets. Jacqueline A. Patel (Noise and Vib. Section, Health and Safety Lab., Harpur Hill, Buxton, Derbyshire SK17 9JN, UK, jacqueline.patel@hsl.gov.uk)

European Standard EN 271: 1995 deals with the construction of air-fed shotblasting helmets. It specifies minimum requirements for face, eye, and respiratory protection and tests the noise generated by the breathing air supply. However, it does not test the helmet's ability to attenuate the very high levels of noise generated by the shotblasting process. This is a major omission. A test method was developed for measuring the noise attenuation provided by shotblasting helmets against typical shotblasting noise. It involved objective measurements using the head and torso simulator (HATS) currently prescribed by EN 271, fitted with miniature microphones. Using such a HATS was shown to give attenuation values that correlate well with those measured using human subjects. Results from this project showed: the HATS already used by manufacturers to show compliance of their product with EN 271 can also be used to provide information on the helmet's noise attenuation; and the same HATS can be used, in place of human subjects, to measure the air supply noise according to the method defined in EN 271. The findings from this work will be used to try to influence future revisions of EN 271 so that the noise attenuation of shotblasting helmets is considered.

3:15–3:30 Break

3:30

4pNS6. Effect of active noise reduction (ANR) and passive hearing protectors on detectability of a vehicle backup alarm in noise. John G. Casali (Auditory Systems Lab., Virginia Tech, Blacksburg, VA 24060, jcasali@vt.edu), Erika Dabney (Xerox Corp., Rochester, NY), Gary S. Robinson (Virginia Tech, Blacksburg, VA 24060), and Dan Gauger (Bose Corp., The Mountain, MA)

In a within-subject laboratory experiment, ten subjects of normal-hearing abilities listened for a standard vehicle backup (reverse) alarm that was presented against either an 85 or 100 dBA background noise having either a pink or red spectrum. Subjects' masked thresholds to the alarms were determined under four listening conditions: ANR earmuff, passive earmuff, foam earplug, and unprotected (at 85 dBA only). To become immersed in a workload situation, subjects continuously attended to a computer-generated visual monitoring task while performing the auditory detection experiment. Analysis of variance and *post-hoc* tests were applied to the mean masked threshold results. In 85 dBA noise, masked thresholds were significantly lower when any protector was worn, as compared to the unoccluded condition, lending support for the use of hearing protection with normal hearers when backup alarm signals need to be heard in noise. The ANR earmuff provided a significant advantage over the passive earmuff in both levels of the red noise, and for both noise spectra at 100 dBA. However, the foam earplug yielded lower masked thresholds than either other protector in 100 dBA noise. Subjects' perceptions about efficacy of warning signal detection and about comfort revealed no differences among the three protectors.

3:50

4pNS7. Hearing protector performance for impulse noise. Daniel L. Johnson (Brüel Bertrand & Johnson Acoustics, 4719 Mile High Dr., Provo, UT 84604)

Using Army volunteers, a series of studies were performed at the Blast Overpressure Test Site in New Mexico to establish safe exposure levels for impulse noise. The results of these studies demonstrated that for properly worn hearing protection, the limit of human exposure is set by the threshold of nonauditory damage. This nonauditory threshold can be depicted by a simple formula based on peak pressure, *A*-duration, and number of impulses. Using this information, a standard for impulse noise will be proposed in which the attenuation of the hearing protective need not be measured or predicted. The protector need only to be shown to protect adequately. One method for evaluating this adequacy is based on a procedure that verifies that Temporary Threshold Shifts of Hearing are not occurring. Another method is to measure the total A-weighted energy under the protection device or devices. The standard will also include a computer model that will provide the means for the user to predict the risk of hearing injury by analyzing the waveform of the impulse.

Contributed Papers

4:10

4pNS8. A statistical classifier for hearing protector REAT data. William J. Murphy and John R. Franks (Hearing Loss Prevention Section, NIOSH, M.S., C-27, 4676 Columbia Pkwy., Cincinnati, OH 45226-1998)

Hearing protector attenuation data have been demonstrated to exhibit bimodal characteristics due to poorly fitting protectors. The data have been modeled with two distributions [Murphy and Franks, J. Acoust. Soc. Am. **106**, 2262 (1999)]. The bimodal distribution model was used as the basis of a statistical classifier for an individual's attenuation data to identify whether or not the data come from a poor or well fit distribution. Training classes were developed from ANSI S12.6 Method B subject fit data [Franks *et al.*, Ear Hear. **21**, 218–226 (2000)] for the EAR Classic ($n=25$), EAR Express ($n=20$) and Howard Leight MAX ($n=25$) earplugs. Real ear attenuation at threshold was measured under circumaural headphones for each earplug with a panel of 20 subjects. Preliminary results suggest that data from 500 or 1000 Hz yield the best ability to separate the attenuations.

4:25

4pNS9. The measurement of impulse noise protection by earmuffs with the use of various sound sources. Jan Zera (Central Inst. for Labour Protection, ul. Czerniakowska 16, 00-701 Warsaw, Poland)

Impulse attenuation of 23 hearing protectors from various manufacturers was determined by the transmission-loss method. Impulses were generated by a foil blaster, a starter gun, or by hitting a steel plate with a hammer. Peak levels of the impulses ranged from 130 to 160 dB. The measurements were carried out using an artificial test fixture (ISO/TR 4869-3:1989), and a KEMAR manikin. Peak sound pressure level of acoustic impulses was measured with two microphones, one of which was placed under the earmuff and the other one outside the earmuff. The best agreement between impulse attenuation and the protectors' M, L, and SNR parameters was found when the foil blaster was used as a sound source. The results also show that the attenuation measured for impulses is usually smaller than the H, M, L, and SNR values given in the product data specification. [Work supported by the Polish Committee for Scientific Research Grant No. SPR 03.7.21.]

4:40

4pNS10. Noise attenuation of an insertion-type hearing protector by mathematical model and experimental verification. Samir N. Y. Gerges, Elizabete Y. N. Bavastri (Federal Univ. of Santa Catarina (UFSC), Mech. Eng. (EMC-LVA), Trindade, Florianopolis, SC, Brazil, CEP: 88040-9000, gerges@mbox1.ufsc.br), and Roberto A. Tenenbaum (COPPE, Univ. Federal de Rio de Janeiro, Rio de Janeiro, Brazil)

The experimental real ear attenuation at threshold (REAT) method is based on the standards ANSI S12.6-1997 and ISO 4869-1990 which consider measurements on qualified subjects, in a qualified acoustic chamber. The REAT method is very expensive, time consuming, and has a very poor frequency resolution. Mathematical and numerical models, such as the one-dimensional plan wave model, finite-element method, and boundary-element method offer cheap, quick, and high-resolution results, especially for comparison and product development. In this paper, the outer ear canal is considered as a straight tube with an open end and the other end closed with an eardrum membrane of human acoustic impedance. For the open ear canal case (without hearing protector on), plan acoustic waves are considered as incident and reflected waves. In the case of a blocked ear (with the insertion hearing protector on), a model of three mediums is considered; air hearing protector air. In this model only one longitudinal wave propagation is considered. Material density and speed of sound are the only parameters considered. The noise insertion loss is calculated as the difference of the sound pressure level at the eardrum without and with the protector on. The results are compared with controlled experimental measurements.

Session 4pPA

Physical Acoustics, Noise, Standards Committee S2, Signal Processing in Acoustics and Structural Acoustics and Vibration: Novel Optical Techniques for Measuring Surface Vibration

James M. Sabatier, Cochair
National Center for Physical Acoustics, University of Mississippi, Coliseum Drive, University, Mississippi 38677

Benjamin A. Bard, Cochair
19 Benjamin Road, Belmont, Massachusetts 02478

Chair's Introduction—1:25

Invited Papers

1:30

4pPA1. Whole field vibration measurements with laser-based techniques. Amit K. Lal and Cecil F. Hess (MetroLaser, 18010 Skypark Circle, Ste. 100, Irvine, CA 92614)

This paper discusses two techniques, a time-domain technique and a space-domain technique, to remotely measure 2-D vibration characteristics of a surface. The first technique is based on using multiple diode laser vibrometers to simultaneously measure the surface vibration at multiple points. The principles of laser Doppler vibrometry are briefly discussed and various architectures of systems to perform landmine detection are presented and discussed. Surface velocity data corresponding to various types of soils shaken by a buried geophone are presented. The second technique is based on laser speckle interferometry using a single digital camera to instantaneously capture quantitative phase information. The phase information of two consecutive frames is then converted to velocity profiles of the vibrating surface. The principles of a new method to perform instantaneous phase shifting interferometry are presented and modal analysis data of turbine blades obtained with an all-digital optical system are discussed.

2:00

4pPA2. Vibration analysis by digital shearography. W. Steinchen, G. Kupfer, and P. Mäckel (FB 15, Dept. of Mech. Eng., Lab. of Photoelasticity, Holography and Shearography, Univ. of Kassel, D-34109 Kassel, Germany)

Digital shearography, a laser interferometric technique in conjunction with the digital image processing, has the potential for vibration analysis due to its simple optical system and sensitivity against small rigid body motions. This paper will focus on its recent developments for vibration analysis and for nondestructive testing (NDT) by dynamic (harmonic) excitation. With the introduction of real-time observation using an automatically and permanently refreshing reference frame, both small and large rigid body motions are greatly suppressed. The development of a smaller and more mobile measuring device in conjunction with a user-guided comfortable program SHEARWIN enables the digital shearography to be applied easily as an industrial on-line testing tool.

2:30–2:45 Break

Contributed Papers

2:45

4pPA3. Laser-ultrasonic measurements of thin-film properties using surface acoustic waves. Donna C. Hurley (Natl. Inst. of Standards & Technol., Boulder, CO 80305, hurley@boulder.nist.gov) and Adrian J. Richards (CSIRO Telecommunications & Industrial Physics, Lindfield, NSW 2070 Australia)

High-frequency, laser-ultrasonic methods to evaluate the mechanical properties of thin films have been developed. The approach is based on the optical generation and detection of surface acoustic waves (SAWs) to determine the frequency dependence of the phase velocity (dispersion relation). Broadband or variable-frequency, quasi-plane wave SAWs are generated by a line-focused, 200-ps pulsed laser. The out-of-plane SAW displacement amplitudes are measured using a path-stablized Michelson interferometer with line-focus detection. The apparatus incorporates differential photodiode detection with a -3-dB bandwidth of 40 kHz–800 MHz. With this system, several displacement waveforms over a range of propagation distances are acquired and used to determine the SAW dispersion. Dispersion relations up to 200 MHz are presented for a variety of specimens, including a series of Si wafers with TiN films 0.2–1.3 micrometers thick and containing residual stress. The results illustrate the applicability of this technique to various conditions such as stiffening and loading films and isotropic and anisotropic substrates. The experimental dispersion curves obtained with this method are compared to theoretical predictions to evaluate the film elastic properties.

3:00

4pPA4. Optical matched-filter and phased-array detection of surface acoustic waves. Todd Murray and Sridhar Krishnaswamy (Ctr. for Quality Eng., Northwestern Univ., Evanston, IL 60208-3020)

A method to optically measure surface acoustic displacements simultaneously over an array of detection points is described. Optical phase gratings are used to create a detection array of laser beams that is directed to the specimen. The scattered beams from the detection array are then collected and combined with a single reference beam in a photorefractive crystal in a multiplexed two-wave mixing configuration. The output beams of the resulting multiplexed photorefractive interferometer are imaged onto either a single photodetector (matched-filter scheme) or to separate

photodetector elements (phased-array scheme). The optical matched filter requires less electronic overhead. By using different object beam footprints, it is possible to configure optical matched-filter receivers that selectively detect SAWs (broadband, narrow-band, chirped, etc.) with enhanced sensitivity. The responses of equispaced and chirped array receivers to propagating SAWs are discussed. The more general phased-array configuration, although requiring greater electronic overhead, provides time-domain signals from each element for postprocessing. This system can therefore be synthetically focused by applying appropriate phase shifts between the detected signals to create images of the acoustic scattering region. Both these multiplexed optical detection schemes, with their high spatial resolution and subnanometer displacement sensitivities, have applications in nondestructive testing and materials characterization.

3:15

4pPA5. Measurements of wineglass resonance using a fiber-optic probe. Andrew A. Piacsek (Dept. of Phys., Central Washington Univ., 400 E. 8th Ave., Ellensburg, WA 98926)

A fiber-optic probe (MTI-2000 Fotonic Sensor) is used by Central Washington University physics students to measure the peak displacement of such vibrating surfaces as acoustically driven wineglasses, parametrically shaken liquids, and soap films. This presentation will consist of a brief description of the fiber-optic measurement system, a discussion of its advantages and limitations, and sampling of some recent data obtained by students.

3:30

4pPA6. Phase-amplitude-modulated laser and microwave vibrometers. Dimitri Donskoy, Nikolay Sedunov, and Edward Whittaker (Stevens Inst. of Technol., Hoboken, NJ 07030)

The paper presents results of a new development of remote vibration sensors using modulated electromagnetic (optical and microwave) signals. The microwave sensor (vibrometer) utilizes the effect of phase modulation of a cw microwave carrier reflected from vibrating surfaces or subsurfaces. The optical vibrometer uses a similar principle of phase modulation of the microwave frequency modulating envelope of the optical carrier. Both optical and microwave vibrometers can share the same signal, generating and receiving electronics and algorithms. Using different carriers but shared electronics, both sensors can be combined in one device offering versatile capabilities for various applications. The developed concept has been experimentally validated for both optical and microwave sensors. The 2.5-GHz portable vibrometer has been designed and tested. The description of the vibrometers, examples of their applications, and comparison with commercial LDV systems will be discussed in the presentation.

3:45

4pPA7. Measurement of grain size using diffused ultrasound energy. Guy Lamouche, Martin Lord, and Andr Moreau (Industrial Mater. Inst.–NRC, 75 de Mortagne, Boucherville, QC J4B 6Y4, Canada, guy.lamouche@nrc.ca)

A new method is proposed to determine grain size in steel sheets by combining a laser-ultrasound technique with ultrasound energy diffusion. A short laser pulse impinging on a metal surface generates an ultrasonic wave packet that propagates through the material. This wave packet is scattered by the discontinuities it encounters at the boundaries of the metallic grains. At the beginning, this energy is quite localized near the generation point and then spreads in the sheet in a way that can be described as a two-dimensional diffusion process. This diffusion process can be followed by measuring the ultrasound field amplitude at the surface of the sheet at various distances from the generation point. This observation is performed remotely by using a laser light that is frequency shifted by the surface movement (Doppler effect). Theoretical considerations suggest that there should exist a relationship between the size of the grains responsible for the ultrasound scattering and the diffusion coefficient describing the ultrasound energy diffusion process. We present preliminary experimental results showing that such a relationship does exist, thus providing a method to remotely measure grain size in steel sheets.

4:00

4pPA8. The modes and loss mechanisms of a high *Q* mechanical oscillator. Xiao Liu, J. F. Vignola (SFA, Inc., Largo, MD 20774), S. F. Morse, D. M. Photiadis, A. Sarkissian, M. H. Marcus, and B. H. Houston (Naval Res. Lab., Washington, DC 20375)

In order to understand the unusually high Q nature of the double paddle oscillator employed in a number of recent experiments, laser Doppler vibrometry (LDV) measurements of the normal modes of the oscillator were carried out. The observed resonance frequencies and mode shapes are in excellent agreement with three-dimensional finite element simulations. The second antisymmetric torsional mode was found to have exceptionally good vibration isolation because of its mode shape. This explains its extremely small low temperature ($T<10$ K) internal friction (2×10^{-8}). A design criterion is established to minimize the damping of similar oscillators by correlating features of the measured and predicted mode shapes.

Session 4pSA

Structural Acoustics and Vibration: Sound/Structure Interaction

Karl N. Grosh, Chair

Department of Mechanical Engineering and Applied Mechanics, University of Michigan, 2350 Hayward Street, Ann Arbor, Michigan 48109-2125

Contributed Papers

1:30

4pSA1. Computational evaluation of satellite equipment panel modal densities and radiation efficiencies. Stephen C. Conlon (Hughes Space and Communications Co., Bldg. S24 MS D563, P.O. Box 92919, Los Angeles, CA 90009) and Stephen A. Hambric (Appl. Res. Lab., Penn State Univ., State College, PA 16804)

During the design phases of communications satellites Statistical Energy Analysis (SEA) methods are often used to predict critical equipment panel structural-acoustic vibration responses. Difficulties arise in the SEA modeling due to the nonhomogeneous nature of these complex equipment panels. Panel constructions are typically sandwich honeycomb (thick) cored panels with internal embedded heat pipes, external attached stiffeners, external thermal doubler plates, as well as the attached equipment itself. The resulting built up structure is very difficult to analyze using SEA theoretical asymptotic values. This work computationally evaluates the modal density and radiation efficiency of a typical communications satellite equipment panel utilizing a detailed finite-element model along with a lumped parameter radiation representation developed by Fahnline (1997). The full built up nonhomogeneous panel, as well as the uniform baseline panel, are evaluated using both idealized and realistic boundary conditions. The characteristics of the nonhomogeneous equipment panels are shown to vary significantly from commonly used SEA values. Simple curve fits are developed for future SEA evaluations of similar structures.

1:45

4pSA2. Far-field radiation from a source in a flat rigid baffle of finite size. Joerg W. Panzer (New Transducers Ltd., Huntingdon PE29 6FW, UK, j.panzer@nxtsound.com)

The far-field radiation of a source in a flat rigid baffle of finite size is derived. The velocity distribution of the source is arbitrary as well as the shape of the source and the baffle. The derivation covers the two-sided as well as the single-sided radiating case. The approach makes use of the Kirchhoff–Helmholtz equation as well as of spatial Fourier transforms of the source and the baffle shape.

2:00

4pSA3. An analytical solution for the self and mutual radiation resistances of a rectangular plate. W. L. Li (United Technologies Carrier Corp., A+R Bldg., Carrier Pkwy., Syracuse, NY 13221)

It is widely accepted that, for an arbitrary acoustic wave number, the radiation resistance of a simply supported rectangular plate has to be calculated through numerical integration. In this study, an analytical solution for the self and mutual radiation resistances is obtained in the form of the power series of the nondimensional acoustic wave number. Unlike the previous analytical or asymptotic solutions, it is not subject to any of the restrictions usually imposed upon the acoustic wave number. A few numerical examples are given simply to verify the solution and elucidate the reliance of the self and mutual radiation resistances on the acoustic wave number and the plate aspect ratio. It is shown that the present formulae are extremely efficient by cutting the CPU time by orders of magnitude in comparison with the traditional numerical integration scheme. This investigation successfully fills the long-existing gap in solutions for the moderate wave numbers that are often of primary concerns in an acoustic analysis.

2:15

4pSA4. Introduction of complex layer analysis (CLA) for acoustic radiation problems. David Feit and Daniel T. DiPerna (CDNSWC, 9500 MacArthur Blvd., W. Bethesda, MD 20817, feitd@nswccd.navy.mil)

The pressure radiated by a planar vibrating surface with a prescribed harmonic time-varying surface velocity can be represented as the inverse Fourier transform of the product of the velocity transform and the spectral form of the acoustic surface impedance. This integral representation usually cannot be evaluated in closed form, but can, however, be evaluated asymptotically in the far field, i.e., at field points large compared to the acoustic wavelength. In this note a novel approach, ''complex layer analysis'' (CLA) is introduced which involves an approximation of the acoustic impedance in terms of a rational function. Contour integration is then used to obtain an explicit expression for the near-field pressure distribution as a sum of waves propagating with wave numbers corresponding to the poles of the approximate acoustic impedance. Such an expression makes explicit the notion of pressure-field generation localizing to points of discontinuity in the velocity distribution. [Work supported by the Carderock Division Naval Surface Warfare Center ILIR Program.]

2:30

4pSA5. Determination of design trends in the structural acoustic optimization of a multi-ply laminated composite cylindrical shell. W. M. Johnson and K. A. Cunefare (School of Mech. Eng., Georgia Tech, Atlanta, GA 30332-0405)

Structural acoustic optimization (SAO) is a multidisciplinary design technique used to determine the optimal acoustic properties of a structural design. Previous works have applied structural acoustic optimization to stiffened and unstiffened aluminum cylindrical shells subject to acoustic and structural vibration excitation. Most recently, SAO has also been applied to multiply laminated composite cylinders. In this case the ply angles served as the design variables but were not allowed to vary over the cylinder length. In this work we perform a structural acoustic optimization on a laminated composite cylinder with variable ply angles along the cylinder length. This is accomplished with a computational design tool incorporating finite- and boundary-element methods. It is anticipated that these results will help to clarify what design trends control the interior acoustic environment. [This work was supported by the Structural Acoustics Branch of NASA Langley Research Center and Graduate Student Research Program Grant (NGT-1-52167).]

2:45

4pSA6. Predictive approaches and measurements of far-field radiation from submerged partially coated (compliant layer) cylindrical shells. Peter C. Herdic, Brian H. Houston, Douglas M. Photiadis, and Earl G. Williams (Naval Res. Lab., 4555 Overlook Ave., Washington, DC 20375, herdic@code7136.nrl.navy.mil)

In-water laboratory experiments involving spatially dense near-field and far-field responses ($ka=1$ to 20) are reported for a number of point driven cylindrical shells with varying amounts of compliant material attached to the exterior. The first framed shell has a square opening in the otherwise fully coated surface. Two modeling approaches are employed to predict the far-field radiation from this opening. The first method is based on the Helmholtz Integral Equation (HIE), using a simple baffled plate model where only surface quantities around the opening are projected. A square piston in a cylindrical baffle is also implemented as a second predictive approach. The predictions are found to be in general agreement with the actual radiated values determined through holographic projections to recover the off-board quantities. The discussion includes the conditions necessary for an accurate model. Additionally, two shells with different levels of interior structural complexity are each investigated using near-field acoustic holography (NAH) where the degree of compliant material coverage is varied. The physics of these systems will be discussed, including the nature of the normal modes, the energetics, and the degree to which interesting phenomena such as ''tunneling'' are observed.

3:00–3:15 Break

3:15

4pSA7. Numerical and experimental characterization of the transmission loss of complex composite panels. Maxime Bolduc, Raymond Panneton, Noureddine Atalla, and Jean-Luc Wojtowicki (GAUS, Mech. Eng. Dept., Univ. de Sherbrooke, 2500 boul de l'universit, Sherbrooke, QC J1K 2R1, Canada)

The noise reduction in aircrafts is one of the top priorities for the aerospace industry. In this paper, a thorough investigation is done on the reliability of numerical FEM/BEM methods to predict the sound transmission loss of aircraft-type panels. The investigation is based on the comparison between numerical predictions, using commercial and in-house software, and measurements for different configurations of panels: single plate, double plate with air cavity, and double plate with cavity absorption. Clamped boundary conditions are used for both the simulation and experiments. The porous material is modeled following Biot theory and the diffuse sound field is simulated as a superposition of plane waves at various directions. For the experimental part, the transmission loss is obtained by setting the panel between a reverberant and an anechoic room. The transmitted power is measured with an intensity probe allowing fine characterization of the effects of complexities such as the boundary conditions and leaks. Furthermore, the quadratic velocity is measured as a means of validation for the mode shapes. Discussions on the numerical procedure to follow for providing reliable transmission loss predictions and a discussion on the discrepancies between the experiments and the numerical simulations will be done.

3:30

4pSA8. Transmission loss of heterogeneous porous materials. Noureddine Atalla (GAUS, Dept. of Mech. Eng., Univ. de Sherbrooke, QC, Canada) and Franck Sgard (LASH, DGCB URA CNRS 1652, ENTPE, 69518 Vaulx-en-Velin Cedex, France)

Recently, the authors have investigated the acoustic absorption performance of heterogeneous materials made up from a combination of thin porous patches, with different acoustic properties. It has been shown that absorption could be significantly increased in the case of properly designed macroperforated wools in a large frequency band. This paper investigates the transmission loss of similar macroheterogeneous materials. The studied materials are made up of porous elastic materials containing solid inclusions. They are placed into an infinite rectangular waveguide. The numerical model used combines a finite-element formulation for the heterogeneous porous material and a modal description of the acoustic field in the waveguide. The coupling between the solid and the porous patches is accounted for naturally through the use of a modified Biot's formulation. Numerical results are presented to show the effect of several geometric and physical parameters on the transmission loss performance of heterogeneous porous materials.

3:45

4pSA9. A new finite/infinite scheme to model acoustic trim under realistic working conditions. Jean-Pierre Coyette, Gregory Lielens, and Jean-Louis Migeot (Free Field Technologies, 16 Pl. de l'Universit, B-1348 Louvain-la-Neuve, Belgium)

The paper presents a finite-element scheme modeling all aspects of the behavior of acoustic trim: insulation, absorption, and damping. The models can combine different materials (viscoelastic, poroelastic, incompressible), acoustic fluid, and include complex effects like visco-thermal dissipation in thin air layers. Unbounded acoustic regions are modeled using stabilized conjugated infinite elements that present exceptional convergence properties. The structure can be subjected to a number of different excitations: kinematical (fixed motion), dynamical (discrete or distributed forces) or acoustical. An incident diffuse sound field or distributed pressure fluctuations corresponding to aerodynamic boundary layers can also excite the structure. The presentation will be organized as follows: (1) a brief review of the techniques involved: finite-element method for the structure and the acoustic near field, infinite elements for the acoustic far field; (2) material models considered: viscoelastic, poroelastic, incompressible, acoustic (including viscothermal effects); (3) excitation models: kinematical, dynamical, acoustical, aerodynamic; (4) validation of the approach; (5) three application cases: comparison of the dynamic behavior of a treated and untreated shell, acoustic transparency of an insulating sandwich, evaluation of the acoustic impedance of a porous layer.

4:00

4pSA10. Dynamic structural–acoustic–piezoelectric finite-element analysis of a sonar array for a supercavitating high-speed underwater vehicle. Robert M. Koch (Adv. Technol. Div. Naval Undersea Warfare Ctr., 1176 Howell St., Newport, RI 02841-1708, kochrm@npt.nuwc.navy.mil)

Three-dimensional nonlinear dynamic finite-element analysis (FEA) is utilized herein in order to examine the coupled structural, acoustic, and piezoelectric noise propagation physics within a high-frequency sonar array for a supercavitating underwater research vehicle. Specifically, steady-state dynamic simulations of the sonar array's coupled performance are carried out that model both (a) the acoustic radiation field resulting from electrically actuated ceramic array elements, and (b) the ceramic array element electrical response due to incident structure-borne excitation; fluid-borne excitation is being examined presently and may also be presented if completed in time. This paper focuses primarily on the interesting issues related to accurately modeling such a complex coupled system. For example, the frequency-dependent properties of a nonlinear elastomer absorption material within the array must be carefully characterized at very high frequencies in order to gain accurate results. Additionally, determining electrical properties of the piezoelectric ceramic material can pose a significant problem. Finally, performing a three-dimensional simulation at high frequencies with coupled physics analyses is not a trivial undertaking.. Laboratory shaker experiments performed in order to validate the modeling will also be presented.

4:15

4pSA11. Effects of tube shape on nonlinear resonant oscillations of interior gas. Young-Doo Chun and Yang-Hann Kim (NOVIC, Dept. of Mech. Eng., KAIST, Science Town, Taejon 305-701, Korea, yhkim@mail.kaist.ac.kr)

An investigation is conducted to get an outline of designing the tube shape for a sonic compressor with high efficiency. When a closed tube oscillates with its resonant frequency, the interior acoustic properties such as density, velocity, and pressure undergo very large perturbation with

nonlinear patterns. In order to analyze the nonlinear acoustic phenomena such as harmonic generation, shock formation, and resonant frequency shift, two-dimensional axisymmetric governing equations are derived and solved numerically. Numerical simulations are performed to investigate the effect of the tube shape that gives us maximum possible pressure. Results show that the resonant frequency and patterns of pressure waves strongly depend on not only the tube shape but also the amplitude of driving acceleration. The degree of nonlinearity of pressure signals is measured by the newly defined nonlinear energy ratio of the signals. The nonlinear energy ratio well describes not only the energy transfer to higher harmonics but also the compression efficiency, so that it can be a good measure for designing tube shape.

THURSDAY AFTERNOON, 7 DECEMBER 2000 PACIFIC SALONS A, B AND C, 1:25 TO 5:15 P.M.

Session 4pSC

Speech Communication: Kenneth Stevens' Contributions to Speech Research: Influences and Future Directions

Joseph S. Perkell, Cochair
Massachusetts Institute of Technology, Room 36-591, 50 Vassar Street, Cambridge, Massachusetts 02139-4307

Stefanie R. Shattuck-Hufnagel, Cochair
Speech Communication Group Research Laboratories, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Chair's Introduction—1:25

Invited Papers

1:30

4pSC1. Ken Stevens and cross-linguistic phonetic studies. Peter Ladefoged (Phonet. Lab., Linguist. Dept., UCLA, Los Angeles, CA 90095-1543, oldfogey@ucla.edu)

Many years ago we realized that one person's pathological voice quality might be a necessary part of another person's phonological contrast. Some American English speakers may have a very breathy voice, but a Gujarati speaker may need that voice quality to distinguish the word meaning ''outside'' from the word meaning ''twelve.'' It was largely Ken Stevens's work, in personal communications, that enabled us to quantify these distinctions. This led to the notion that it was possible to describe a continuum of phonation types: voiceless, breathy, voiced, creaky, glottal closure. Using a modified version of another notion suggested by Stevens, we were able to consider possible quantal states along the parameter of glottal stricture. Further investigation of different languages suggested that contrasts along a single parameter may not be as straightforward as they seemed, and a number of ''helping features'' (yet another Stevens notion) were often involved. Despite being misled by the binary notions of colleagues at MIT, Ken Stevens's work has been of great help in linguistic phonetics.

2:00

4pSC2. A model of speech motor control and supporting data: Influences of quantal effects. Joseph S. Perkell (MIT–RLE, Rm. 36-591, 50 Vassar St., Cambridge, MA 02139) and Frank H. Guenther (Boston Univ., Boston, MA 02215)

A model of speech motor control is presented, along with examples of supporting data. In the model, segmental speech movements result from a process which plans a trajectory that passes through a sequence of phoneme-specific auditory goals. Some auditory goals are constrained by quantal relations between the vocal-tract area function and the resulting acoustics, as Stevens has proposed, and some are defined by quantal relations between motor commands and the area function, called biomechanical saturation effects. Both types of quantal relations make it possible to produce sounds that have relatively stable acoustic properties using somewhat imprecise motor commands. Therefore, quantal effects influence motor control strategies of individual speakers. They also influence sound patterns of languages, in that they help to define sounds such as the vowels /i/ and /a/ that occur frequently among languages of the world. Although there are other influences on speech motor control mechanisms and sound patterns, quantal effects are considered to be one of the defining characteristics of spoken language. [Work supported by NIDCD, NIH.]

2:30

4pSC3. The ''noisy'' speech chain. Abeer Alwan (Dept. of Elec. Eng., UCLA, 405 Hilgard Ave., Los Angeles, CA 90095)

Kenneth Stevens' seminal contributions to the field of acoustic phonetics are far reaching and have important implications for modeling normal and disordered speech production and perception. His studies of the relations between underlying discrete linguistic features and the analog worlds of articulation, sound, and hearing have influenced a generation of speech researchers, including myself. Our studies of speech perception in noise are inspired by the work of Ken and his colleagues on uncovering the acoustic and perceptual correlates of linguistic features. I will report on experiments which examined acoustic and perceptual (in noise) cues for the

place of articulation and voicing features for syllable-initial plosive and fricative consonants. Results are analyzed in terms of voicing, vowel context, manner of articulation, place of articulation, shape of the noise masker, and gender of the talker. One of the interesting findings is that certain attributes which cue a feature acoustically and perceptually in quiet conditions (such as the VOT for the feature voicing) do not correlate well with the perceptual robustness of the feature in noise.

3:00

4pSC4. Theoretical approaches to speech perception. Patricia K. Kuhl (Dept. of Speech and Hearing Sci. and Ctr. for Mind, Brain, and Learning, Univ. of Washington, Seattle, WA 98195)

Theoretical approaches to speech perception have shifted dramatically during the last 50 years. Early theories, such as Alvin Liberman's ''motor Theory,'' focused on the psychological mechanisms underlying speech perception rather than listeners' analysis of the acoustic signal. In the 1970's, Ken Stevens' work on the acoustic cues for phonetic categories produced a very different view, an ''invariance theory,'' that turned theoretical attention to the acoustic signal and the detailed phonetic information that listeners gleaned from its analysis. The last decade has seen another theoretical shift. Strongly influenced by developmental studies, new theories (e.g., Kuhl's native language magnet model, or NLM) combine the sensory and psychological approaches, arguing that listeners both perform a detailed analysis of language input and psychologically represent that information in a complex way that is unique to each language. The presentation will discuss these shifts and the kinds of data now necessary to produce the next theoretical advance. [Work supported by NICHD HD37954 and Human Frontiers RG0159 to PKK.]

3:30–3:45 Break

3:45

4pSC5. Human communication with information systems. James Flanagan (Ctr. for Adv. Information Processing, Rutgers Univ., Piscataway, NJ 08854)

As networked information systems increase in complexity, natural modes of communication assume a greater role in the design of human–machine interfaces. Modalities of sound, sight, and touch are largely favored by humans, with conversational interaction being a central medium for information exchange. The distinguished research career of Professor Kenneth N. Stevens, celebrated in this special session on speech communication, has contributed directly to the basic understanding that supports current capabilities in voice interaction. While as yet primitive in comparison to human versatility, the technologies for sound, in concert with those for sight and touch, can be prudently applied to transcend the functionalities of traditional mouse and keyboard. This report describes a research effort in simultaneous utilization and fusion of sound, sight, and touch signaling in a networked collaborative environment. The conversational interaction depends upon automatic speech recognition, speech synthesis, and ''hands-free'' sound capture. Automatic face recognition, visual gesture, and eyetracking serve the sight modality. A force-feedback tactile glove provides manual gesture and grasp of virtual objects. And, intelligent software agents fuse and manage the multiple modes to develop a context-aware estimate of user intent. An experimental system utilizing these interface modes is applied to crisis management, and is demonstrated by videotape. [Components of this research were supported by National Science Foundation Contract Nos. IRI-96-18854 (STIMULATE) and IIS-98-72995 (KDI), DARPA Contract No. N66001-96-C-8510 (DISCIPLE), the Intel Corporation, Texas Instruments, and the Rutgers Center for Advanced Information Processing.]

4:15

4pSC6. Conversational Interfaces: Today and Tomorrow. Victor Zue (MIT Lab. for Computer Sci., 545 Technology Square, Cambridge, MA 02139)

For more than a decade, the Spoken Language Systems Group at the MIT Laboratory for Computer Science (with members made up of four generations of Ken Steven's disciples), have been conducting research leading to the development of conversational interfaces that enable users to access and manage information using spoken dialogue. Our research has been heavily influenced by Ken's teaching and research, ranging from acoustic landmark detection, acoustic-phonetic modeling, hierarchical phonological modeling using subword units such as syllables and features, lexical access strategies, and speech synthesis. While we are far from being able to develop systems capable of carrying on unrestricted conversations with users, we are beginning to field systems that have restricted domain expertise. For example, users can now speak to computers over the telephone, and receive weather, flight arrival–departure, and traffic information. One can even conduct air travel planning, and receive flight and fare information, all by simply speaking to the computer. This talk will highlight some of our recent work, and speculate on the future. [Work supported by DARPA and NSF.]

4:45

4pSC7. Retrieving distinctive features and segments from variable acoustic cues. Kenneth N. Stevens (Res. Lab of Elec. and Dept. of EECS, MIT, 50 Vassar St., Cambridge, MA 02139) and Samuel J. Keyser (Dept. of Ling., MIT, Cambridge, MA 02139)

It is assumed that words are stored in memory as sequences of bundles of distinctive features. Accessing words from the acoustic signal requires that these features be identified and organized into segments. Variability is a major obstacle to this process. Accounting for this variability requires a clear understanding of how variability arises when an utterance is generated by a speaker. It is proposed that variability arises from the operation of two principles: (1) Conservation: some of the defining gestures for a feature in a given segment may overlap with gestures from adjacent segments, causing weakening or masking of certain acoustic cues. (2) Enhancement: depending on the segmental and prosodic environment in which a feature occurs, additional gestures not directly specified by the

feature may be introduced to enhance its perceptual salience. Conservation and enhancement provide a basis on which to select acoustic cues that reflect the gestures that produced them. Selection of acoustic cues for the features of consonant place, voicing, and nasality will be described and it will be shown that combinations of these cues can lead to feature identification with low error rate. [Work supported in part by grants from NIDCD.]

THURSDAY AFTERNOON, 7 DECEMBER 2000 PACIFIC SALONS E AND F, 12:55 TO 5:15 P.M.

Session 4pSP

Signal Processing in Acoustics, Underwater Acoustics and Speech Communication: Blind Deconvolution and Source Separation in Acoustics

Leon H. Sibul, Chair
Acoustics Department, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

Chair's Introduction—12:55

Invited Papers

1:00

4pSP1. Overview of blind signal separation, part I: Criteria and algorithms. Scott Douglas (Dept. of Elec. Eng., Southern Methodist Univ., Dallas, TX 75275)

In blind signal separation, multiple independent source signals are separated from multiple linear mixtures of these signals without specific knowledge of either the source signal characteristics or the mixing conditions. This talk provides an introduction to the blind signal separation task. Three different problem formulations—signal separation of instantaneous mixtures, signal separation of convolutive mixtures, and multichannel blind deconvolution—are described, and their similarities and differences are highlighted. An overview of both information-theoretic and contrast-based separation criteria is then given. Natural gradient optimization procedures, when combined with such criteria, yield simple and useful blind signal separation algorithms. An example of speech separation of real room recordings illustrates the capabilities and limitations of one such approach.

1:30

4pSP2. Overview of blind signal separation, part II: Structures and applications. Kari Torkkola (Motorola Labs, 2100 E. Elliot Rd., MD EL 508, Tempe, AZ 85284)

The focus of this talk is a class of blind signal separation procedures, blind separation of convolutive mixing of signals. The emphasis is on acoustic signals and methods applicable thereof, and their capabilities and limitations. At least two possible taxonomies of separation methods are possible. These are based on how to parametrize the separating structure, and on the criteria used for separation. The former will be the emphasis of the presentation. Different criteria to solve the problem are only briefly touched upon. Possible application areas are enumerated, especially in acoustics. Two exemplary areas are audio signal enhancement in the presence of disturbing noise, and machine vibration analysis. Some issues on why it is not straightforward to move blind separation techniques into real world applications are elaborated. Limitations in the audio setup are discussed that make unconstrained BSS an extremely hard (if indeed solvable) problem. Some further questions are posed and directions for further research are presented.

2:00

4pSP3. Realistic application of acoustic blind source separation. Lucas Parra (Sarnoff Corp., CN5300, Princeton, NJ 08543)

Blind source separation in an acoustic environment is a challenging problem. The reverberation times in realistic rooms require long separation filters. They have to be estimated in very short times as the room responses change drastically with user location. In addition, convolutive blind separation has a number of indeterminations that make the identification of filter coefficients difficult. Decorrelation of changing cross-power spectra of nonstationary signal has been proposed as a separation criterion by various authors. We have presented in particular a frequency domain on-line algorithm that seeks to decorrelate cross-power spectra estimated over multiple times. We will report the performance of the algorithm in a realistic environment. The results will be compared to more conventional adaptive beamforming algorithms. Furthermore, speech recognition results for a command and control application with a distant talking user will be presented.

2:30

4pSP4. Blind source separation and blind deconvolution in experimental acoustics. Michael J. Roan and Josh Erling (The Appl. Res. Lab., Penn State Univ., P.O. Box 30, University Park, PA 16802, mjr110@psu.edu)

In active noise control, vibration control, and noise analysis, observed signals are typically a mixture of many sources, possible convolved with an unknown filtering function. A new area of research in the area of blind source separation and blind devonvolution (which has had application on the area of wireless communications) has promising application in noise and vibration control,

radar/sonar, and noise analysis. The area of research exploits information theoretic criteria (rather than higher statistics) for blind source separation (BSS) and blind deconvolution (BDC). In this paper, several algorithms for BSS and BDC are presented. These algorithms are applied to experimental data. Experimental work in the areas of BSS has been with the separation speech signals from several speakers, and BDC work has been with speech signals convolved with artificial filters. It is shown through experimental results that several real acoustic filters (i.e., driven duct, driven duct with expansion chamber, and driven duct with Helmholtz resonator sidebranch) can be deconvolved. It is shown that application of the learned inverse filtering functions to the measured time series removes the effect of the filtering function. Finally methods for applying the algorithms for noise source identification, noise/vibration control, and noise analysis are presented.

3:00

4pSP5. Deriving optimal codes for speech and natural sounds using generalized independent component analysis. Michael S. Lewicki (Computer Sci. Dept. and Ctr. for the Neural Basis of Cognition, Carnegie Mellon Univ., 4400 Fifth Ave., Pittsburgh, PA 15213)

A Bayesian method for inferring an optimal basis is applied to the problem of finding efficient codes for natural sounds. The key to the algorithm is multivariate non-Gaussian density estimation, which is an equivalent independent component analysis when the form of the marginal density is fixed. An important advantage of the probabilistic framework is that it provides a method for comparing the coding efficiency of different bases objectively, and the derived codes can be shown to have better coding efficiency compared to traditional Fourier and wavelet bases. It also provides a method for Bayesian signal denoising and filling in of missing samples using a basis that is optimized to the structures in the data. When this framework is applied to deriving efficient codes of speech and natural sounds, the codes share many of the coding properties of the cochlear nerve. Time–frequency analysis shows that it is possible to derive both Fourier-like and wavelet-like representations by deriving efficient codes for different classes of natural sounds.

3:30–3:45 Break

Contributed Papers

3:45

4pSP6. Blind source separation of mixtures of speech signals with unknown propagation delays. Phillip De Leon and Yunsheng Ma (Klipsch School of Elec. & Computer Eng., New Mexico State Univ., Box 30001, Dept. 3-0, Las Cruces, NM 88003-8001, pdeleon@nmsu.edu)

Blind source separation of mixtures of speech signals has received considerable attention in the research community over the last 2 years. One computationally efficient method employs a gradient search algorithm to maximize the kurtosis of the outputs thereby achieving separation of the source signals. While this method has reported excellent separation results (30–50-dB SIR), it assumes a simple linear mixing model. In the general case, convolutional mixing models are used, however, this is a rather difficult problem due to causality and stability restrictions on the inverse not to mention length requirements in the FIR approximation. Research results with the general problem are modest at best. In this paper, we extend the kurtosis maximization ideas for source separation to include delays in the mixing model to at least account for propagation delays from speakers to microphones. The algorithm is designed to first estimate the relative delays of the sources within each mixture using a standard autocorrelation technique. These delay estimates are then used in the kurtosis maximization algorithm where the separation matrix is now modified to include these delays. Simulation results (using the TIMIT speech corpus) generally indicate good separation quality (10–20 dB) with little additional computational overhead.

4:00

4pSP7. Blind separation of L sources from M mixtures of speech signals. Phillip De Leon and Yunsheng Ma (Klipsch School of Elec. & Computer Eng., New Mexico State Univ., Box 30001, Dept. 3-0, Las Cruces, NM 88003-8001, pdeleon@nmsu.edu)

In many real-world applications of blind source separation, the number of mixture signals, L, available for analysis often differs from the number of sources, M, which may be present. In this paper, we extend a successful and efficient kurtosis maximization algorithm used in speech separation of two sources from two linear mixtures for use in problems with arbitrary numbers of sources and mixtures. We examine three cases: underdetermined ($M<L$), critically determined ($M=L$), and overdetermined ($M>L$). In each of these cases, we present simulation results (using the TIMIT speech corpus) and discuss separation matrix initialization issues and observed algorithm limitations. We find that in the critically determined case, the algorithm performs well (20–40 dB SIR) at separating four sources from four mixtures. For the other cases, our results are mixed. In the overdetermined case (two sources, three mixtures), the algorithm performs well (20–40 dB SIR) and we find that the extra mixtures do not result in better SIR measurements. In the underdetermined case (three sources, two mixtures), we are able to separate out at least one source (sometimes two) with the other output signals each containing pairs of the remaining sources.

4:15

4pSP8. Robust speech recognition in adverse environments by separating speech and noise sources using JADE-ICA. Shubha Kadambe (HRL Labs., LLC, 3011 Malibu Canyon Rd., Malibu, CA 90265, skadambe@hrl.com)

Spoken dialogue information retrieval applications are the future trend for mobile users. The typical presence of background noise in mobile environments causes significant reduction in speech recognition accuracy unless the recognizers are trained explicitly for each environment. However, it is not practical to train recognizers for every environment. One solution to this problem is to use an array of microphones and then separate the speech signals from noise by applying Blind Source Separation (BSS) techniques. BSS based on Independent Component Analysis (ICA) is one of the efficient techniques to separate mixed signals and JADE-ICA is one of ICA's variants that is applicable for non-Gaussian signals. The JADE-ICA is applied here, since generally speech and noise are non-Gaussian signals. An array of two microphones was considered and two signal sources: speech and in-vehicle noise were assumed. Recognition accuracy experiments were conducted on 1831 utterances from 1831 speakers after separating the speech signals from noise using the JADE-ICA. The word recognition accuracy for three conditions: clean speech, noisy speech (SNR = 0 dB) and noise separated speech are 91.9%, 45.2% and 91.9%, respectively. This indicates that using the JADE-ICA to separate speech from noise results in the same recognition accuracy as that of clean speech. The details on JADE-ICA and simulations will be provided in the full paper.

4:30

4pSP9. Convolutive blind source separation of speech signals based on amplitude modulation decorrelation. Jörn Anemüller and Birger Kollmeier (AG Medizinische Physik and Graduiertenkolleg Psychoakustik, Universität Oldenburg, 26111 Oldenburg, Germany, ane@uni-oldenburg.de)

The problem of blind separation of a convolutive mixture of speech signals is considered. Signal separation is performed in the frequency domain. The amplitude modulation decorrelation algorithm for convolutive blind source separation of speech is described, the first ideas of which were presented by Anemüller at the joint ASA/EAA/DAGA meeting, Berlin, 1999. The algorithm is based on the fact that the frequency-specific envelopes of speech signals exhibit correlation across different frequencies. This feature can be used to solve the ''permutation problem'' of frequency-domain-based blind source separation algorithms. Furthermore, it leads to separation of good quality since it results in a high number of constraints which must be fulfilled by the unmixed signals. Results for the separation of speech signals are presented for different mixing scenarios, including real-room reverberant mixing.

4:45–5:15

Panel Discussion

THURSDAY AFTERNOON, 7 DECEMBER 2000 YAWL ROOM, 2:00 TO 3:30 P.M.

Meeting of Accredited Standards Committee (ASC) S3 on Bioacoustics

R. F. Burkard, Chair S3
Hearing Research Laboratory, State University. of New York at Buffalo, 215 Parker Hall, Buffalo, New York 14214

J. Franks, Vice Chair S3
Robert A. Taft Laboratories, 4676 Columbia Parkway, Mail Stop C27, Cincinnati, Ohio 45226

P. D. Schomer, Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 43, Acoustics
2117 Robert Drive, Champaign, Illinois 61821

D. D. Reynolds, Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 108/SC4, Human Exposure to Mechanical Vibration and Shock
3939 Briar Crest Court, Las Vegas, Nevada 89120

H. E. von Gierke, Vice Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 43, Acoustics and ISO/TC 108/SC4, Human Exposure to Mechanical Vibration and Shock
1325 Meadow Lane, Yellow Springs, Ohio 45387

V. Nedzelnitsky, U.S. Technical Advisor (TA) for IEC/TC 29, Electroacoustics
National Institute of Standards and Technology (NIST), Sound Building, Room A147, 100 Bureau Drive, Stop 8221, Gaithersburg, Maryland 20899-8221

Accredited Standards Committee S3 on Bioacoustics. The current status of standards under preparation will be discussed. In addition to those topics of interest, including hearing conservation, noise, dosimeters, hearing aids, etc., consideration will be given to new standards which might be needed over the next few years. Open discussion of Committee reports is encouraged.

Scope of S3: Standards, specifications, methods of measurement and test, and terminology in the fields of mechanical shock and physiological acoustics, including aspects of general acoustics, shock, and vibration which pertain to biological safety, tolerance, and comfort.

Meeting of Accredited Standards Committee (ASC) S1 on Acoustics

G. S. K. Wong, Chair S1
Institute for National Measurement Standards (INMS), National Research Council, Ottawa, Ontario K1A 0R6, Canada

P. D. Schomer, Chair, U. S. Technical Advisory Group (TAG) for ISO/TC 43, Acoustics
2117 Robert Drive, Champaign, Illinois 61821

T. J. Kuemmel, Vice Chair S1
Quest Electronics, 510 South Worthington, Oconomowoc, Wisconsin 53066

H. E. von Gierke, Vice Chair, U.S. Technical Advisory Group (TAG) for ISO/TC 43, Acoustics
1325 Meadow Lane, Yellow Springs, Ohio 45387

V. Nedzelnitsky, U.S. Technical Advisor (TA) for IEC/TC 29, Electroacoustics
National Institute of Standards and Technology (NIST), Sound Building, Room A147, 100 Bureau Drive, Stop 8221, Gaithersburg, Maryland 20899-8221

Accredited Standards Committee S1 on Acoustics. Working group chairs will report on their preparation of standards on methods of measurement and testing, and terminology, in physical acoustics, electroacoustics, sonics, ultrasonics, and underwater sound. Work in progress includes measurement of noise sources, noise dosimeters, integrating sound-level meters, and revision and extension of sound level meter specifications. Open discussion of Committee reports is encouraged.

Scope of S1: Standards, specifications, methods of measurement and test and terminology in the field of physical acoustics including architectural acoustics, electroacoustics, sonics and ultrasonics, and underwater sound, but excluding those aspects which pertain to biological safety, tolerance and comfort.

Session 5aAA

Architectural Acoustics: Speech Privacy and Speech Intelligibility in the Built Environment

Kenneth P. Roy, Chair
Innovation Center, Armstrong World Industries, P.O. Box 3511, Lancaster, Pennsylvania 17604

Chair's Introduction—9:00

Invited Papers

9:05

5aAA1. Privacy with intelligibility—A novel electroacoustic solution. Peter A. Mapp (Peter Mapp Assoc., Colchester, Essex C03 4JZ, UK)

The concept of electronic sound masking is a well-established and successful solution in helping to achieve adequate speech privacy either within or between office areas or similar spaces. In many instances, separate paging or voice-alarm systems are also installed. Whereas locating loudspeakers in the ceiling void or within a plenum assists sound diffusion and evenness of coverage, this also degrades speech intelligibility and causes significant nonlinear (frequency-dependent) transmission losses. Conversely, installing conventional loudspeakers directly into the ceiling can result in good intelligibility but poor sound diffusion, leading to uneven sound distribution and immediate localization of the loudspeaker as the sound source. The diffuse, substantially decorrelated and wide dispersion characteristics of the distributed mode loudspeaker suggests that systems employing this new class of device may simultaneously satisfy the requirements for both speech privacy and speech intelligibility. The paper presents the results of a study using a range of loudspeaker formats to establish whether these apparently opposing requirements can be achieved.

9:25

5aAA2. Architectural repair of the acoustics in Koubek Auditorium at the Catholic University of America. Donna A. Ellis (The Catholic Univ. of America, 620, Washington, DC 20064)

Koubek Auditorium is used as the main lecture space in the School of Architecture and Planning at The Catholic University of America in Washington, DC. The space is used for classes in physics, history, and issues in practice, to name a few. It is also regularly used for lectures by professionals from the architectural community. Acoustics in the auditorium are poor and result in ineffective use of the space as well as a poor teaching environment. The main problem areas are the noise levels of the HVAC, reverberation times, and sound distribution. The average noise criteria in this space is NC40. In some locations the maximum is 50–60. At Koubek Auditorium the RT's range from 0.8 s, to 1.46 and higher. This paper will demonstrate how these numbers are brought down to acceptable levels to ensure effective learning in the space. [Work supported by ScantTek, Armstrong, Electro-Media Design.]

9:45

5aAA3. Speech privacy in office spaces. Kenneth P. Roy (Armstrong Innovation Ctr., 2500 Columbia Ave., Lancaster, PA 17603)

Lack of speech privacy in office spaces is an issue that has been inadequately addressed for decades. Now may be the time to reconsider the consequences of not properly designing our work spaces, since ever increasing noise intrusions are having more adverse effects on knowledge worker productivity. Both lab and field tests were performed to characterize the acoustic environment in closed work spaces, and acoustic performance was determined on the basis of the speech privacy index (PI). The goal was to develop an integrated design solution for closed offices spaces based on signal-to-noise performance.

Contributed Papers

10:05

5aAA4. A graphical approach to assessing speech privacy and speech intrusion in offices. David A. Conant (McKay Conant Brook, Inc., 5655 Lindero Cyn Rd., Ste. 325, Westlake Village, CA 91362)

Several very effective closed-form solutions for computing speech intelligibility and speech privacy have been available for designers since Kryters seminal work on articulation index (AI) in the early 1960s. Presented here is a unique graphical solution derived from many of these that lends itself remarkably well to visually assessing, for given room constants and source spectra, the relative contributions of noise insulation class and NC to both speech privacy and intrusion. A principal benefit of this graphical approach is the client's much-improved grasp of the relative merits of current conditions versus alternate designs. Research faculty offices at MIT's new Stata Center, presently in design by architect Frank O. Gehry, are discussed in a case study.

10:20

5aAA5. Differences in performance with different background sound and ambient noise in three open office plans. Heakyung C. Yoon (Carnegie Mellon Univ., Pittsburgh, PA 15213, hcy@andrew.cmu.edu)

The effect of noise on the performance of workers is investigated in three types of open workplaces: a combination of open plan and closed offices, rectangular configuration in open plan, and triangular configuration in open plan. An open workplace is intended to increase workers' performance by creating a knowledge-sharing environment informally and

seamlessly, while it is also known to decrease workers' performance for they are annoyed with or distracted by the sound from the workplace. According to interviews from and questionnaires completed by workers in these workplaces, the order of preference was the combination, the rectangular configuration, and the triangular configuration workplaces. Each subject took two proofreading tests for 10 min under two different sound conditions. One of these conditions was background sound only with quiet HVAC sound and the other was ambient sound, which was recorded from each workplace along with speech and the sound from other activities by a binaural artificial head measurement system. The result is the mean test score of subjects from the triangular configuration workplace is significantly different from the mean test scores of subjects from the combination type workplace and the rectangular configuration workplace, respectively. [Work supported by IBM-Toronto, Inc.]

10:35–10:45 Break

10:45

5aAA6. Speech intelligibility in highway toll booths. Jeffrey P. Feist, Luc Mongeau, and Robert J. Bernhard (Purdue Univ., 1077 Ray W. Herrick Labs., West Lafayette, IN 47907-1077, feistj@ecn.purdue.edu)

The purpose of this study is to improve the working conditions at highway toll collection facilities. It is believed that the sustained noise levels create an unpleasant working condition and hamper communication between attendants and patrons. Controlling noise in this environment is particularly challenging because the toll booth is a partially open enclosure with reflecting walls which may amplify the traffic noise. Toll collection employees were surveyed in order to assess, among other factors, the influence of ambient traffic noise on speech intelligibility. The results suggest that communication problems exist and that active noise control (ANC) headsets are a potential solution. Low-frequency vehicular noises are believed to strongly contribute to the communication problem, and ANC systems are particularly effective at reducing such noise emissions. To adapt ANC headsets to this application, insertion loss measurements have been made for a variety of active headsets. Attention was given to the level of attenuation within the frequency range of speech intelligibility. Modifications are underway to optimize an active headset for speech intelligibility and noise reduction. Additional work includes modeling the acoustics of the environment with a ray-tracing software package. This modeling work will allow structural changes to be evaluated.

11:00

5aAA7. Investigation of speech intelligibility in building with curved surfaces by means of objective measures. Mokhtar Harun, Tharek Abd. Rahman, Md. Najib Ibrahim, and Ahmad Khan Said (Universiti Teknologi Malaysia, 81300 Skudai Johor, Malaysia)

There are two aspects in an acoustics design process that have to be taken care of in order for the room and the sound system to be intelligible. The first is the electronic part of the design. The second is the environment in which the system is to be installed. With the assumptions that the selection of sound system equipment and installation are correct and thorough, the sound reflection pattern and its delay in building with curved surfaces has to be investigated and understood. This paper investigates via measurements as well as through the simulation software the distribution of sound pressure level, sound delay (RT60), time and amplitude of early reflection (R_{early}), and late reflection (R_{late}) in a symmetrical room and in a room with curved surfaces. R_{early} is categorized as the first-arriving wave-front that happens before 35 ms and R_{late} is that within 70 and 120 ms. Preliminary investigation has shown that the equal distribution of these objective measures is destroyed with the existence of curved surfaces in the room. This unequal distribution hints that speech intelligibility of a room with curved surfaces is also varied and altered.

11:15

5aAA8. On the disagreement between speech transmission index (STI) and speech intelligibility. Hiroshi Onaga (Dept. of Architecture, Kinki Univ., Higashi-osaka, Osaka, 577-8502 Japan, onaga@arch.kindai.ac.jp), Yoshihiro Furue (Fukuyama Univ., Fukuyama, 729-0292 Japan), and Tetsuo Ikeda (Kinki Univ., Osaka, 577-8502 Japan)

The change tendency of speech transmission index (STI) was examined by calculation for several modeled sound fields. As a result, it has been revealed that STI contradicts the generally accepted concept of the importance of early energy for speech intelligibility. Namely, for single reflection fields, (1) STI takes the same value for positive and negative levels of reflection relative to direct sound, and for positive and negative delay times. (2) STI takes the lowest value when the reflection level is 0 dB (equal to direct sound), and the increase and decrease of reflection level from 0 dB causes a rise in STI. For multiple reflection fields, (3) STI evaluates only the degree of energy concentration or dispersion in the time domain. Also, energy concentration due to strong reflections with any delay time increases the value of STI. However, the already published experimental results of intelligibility on single reflection fields differ from the above-mentioned properties of STI, (1) and (2). Moreover, the authors' intelligibility experiment with multiple reflections indicates the importance of early energy for intelligibility and shows the disagreement with the properties of STI (3). From the above findings, STI cannot be considered as corresponding to intelligibility.

11:30

5aAA9. Relation between intelligibility and listening difficulty using word lists based on word familiarity. Masayuki Morimoto, Masaaki Kobayashi (Environ. Acoust. Lab., Kobe Univ., Nada, Kobe, 657-8501 Japan, mrmt@kobe-u.ac.jp), and Hiroshi Sato (Tohoku Univ., Aoba, Sendai, 980-8579 Japan)

Listening tests to assess word intelligibility and listening difficulty were performed to obtain a subjective criterion for speech transmission. It is well known that both intelligibility and listening difficulty depend on the choice of test words. The test words used in the present tests were selected according to word familiarity [Sakamoto *et al.*, J. Acoust. Soc. Jpn. **54**, 842–849 (1998)]. The subject was asked to repeat each word and to rate listening difficulty into one of four categories: (0) not difficult, (1) a little difficult, (2) fairly difficult, and (3) extremely difficult. After the tests, the four categories were reclassified into (0) and the others. The listening difficulty is defined as the percentage of the total number of responses except (0). The results demonstrate that: 1) the listening difficulty is highly correlated with intelligibility, 2) the listening difficulty can evaluate the quality of speech transmission more exactly and sensitively than intelligibility, 3) the effects of word familiarity on the listening difficulty and intelligibility do not depend on test sound fields, and 4) STI is not related to either the listening difficulty or intelligibility. Furthermore, the results suggest some other advantages in using listening difficulty as opposed to intelligibility.

11:45

5aAA10. Effects of word familiarity on word intelligibility for listeners with presbycusis in noise and reverberation. Hiroshi Sato, Hiroshi Yoshino, Hayato Sato (Dept. of Archiect. and Bldg. Sci., Grad. School of Eng., Tohoku Univ., Aramaki Aza Aoba, Sendai, 980-8579 Japan), Yoîti Suzuki (2-1-1, Katahira, Aobaku-ku, Sendai, 980-8577 Japan), Shigeki Amano, and Kimihisa Kondo (3-1, Morinosato Wakamiya Atsugi-shi, Kanagawa, 243-0198 Japan)

Listening tests to assess word intelligibility were performed to estimate the effects of presbycusis under control of the word familiarity in experimental sound fields with noise and/or reverberation. The test words used in the present tests were selected to have similar word familiarities [Sakamoto *et al.*, J. Acoust. Soc. Jpn. **54**, 842–849 (1998)]. Moreover, difference/similarity of the effect of word familiarity on speech intelligibility between young listeners with normal hearing and elderly listeners

with impaired hearing was also examined. The results demonstrate that: (1) Comparing the intelligibilities of the elderly with that of the young, the normal hearing subjects can tolerate 5 dB A more intensive noise than impaired listeners. (2) The word intelligibility of young subjects shows a 25% higher score than that of the aged. This result is very similar to sentence intelligibility [Sato *et al.*, Proc. Noise Effects '98 (Sydney, Australia) 168–173 (1998)] and Japanese trisyllable intelligibility [Sato *et al.*, J. Archit. Plann. Environ. Eng. **AIJ**, 15–20 (1997)]. (3) The effect of word familiarity on word intelligibility is almost equal between the young normal and elderly impaired subjects.

FRIDAY MORNING, 8 DECEMBER 2000 SCHOONER/SLOOP ROOMS, 8:10 TO 11:30 A.M.

Session 5aAB

Animal Bioacoustics: William Cummings Session on the Acoustics of Whales and Dolphins II

Whitlow W. L. Au, Chair
Hawaii Institute of Marine Biology, P.O. Box 1106, Kailua, Hawaii 96734

Chair's Introduction—8:10

Contributed Papers

8:15

5aAB1. Comparison of Alaskan and Hawaiian humpback whale song at the song-unit level. Adam S. Frankel (Marine Acoustics, Inc., 901 N. Stuart St. #708, Arlington, VA 22203) and Jan Straley (Univ. of Alaska, Sitka, AK 99835)

Humpback whale songs were recorded in Hawaii and Alaska in 1991, 1994 and 1998. Songs from both areas showed the typical hierarchical theme structure. Strong differences between years were noted, as expected. Songs in Alaska tended to be shorter than those recorded in Hawaii, as previously reported. A unit-level analysis was conducted using an automated measurement program. The program reports >100 descriptive variables. The size of the variable set was reduced using principal components analysis. The resulting principal components (PC) set that accounted for 80% of the variance was tested with discriminant analysis for differences between recordings (nominal individuals) and region. Preliminary analysis indicates that the songs were similar between regions at the theme level. There were no significant differences between the two regions at the unit level of analysis. These results support earlier acoustic and visual-resight data that the two regions are both part of the same population. The small size of the Alaskan sample leaves open the possibility that regional differences exist, but were not detected with this sample.

8:30

5aAB2. Short duration sounds recorded from blue whales (*Balaenoptera musculus*) off Peru. Tom Norris (SAIC Maritime Div., 3990 Old Town Ave., Ste. 105A, San Diego, CA 92110) and Jay Barlow (NMFS Southwest Fisheries Sci. Ctr., La Jolla, CA)

Recordings of low-frequency sounds were made in the presence of blue whales (*Balaenoptera musculus*) in 1998 and 1999 in waters off northern Peru during a marine mammal research cruise. Recordings were made on DAT from sonobuoys deployed near vocalizing whales. Photographic identifications and biopsy samples were collected from several animals that were recorded. Over 20 signals were extracted and analyzed from DAT segments that contained the best signal quality. All signals from the 1998 encounter consisted of rapid, frequency modulated (FM) downsweeps (begin frequency $\bar{x}=94$, S.D.=13.6 Hz; end frequency $\bar{x}=59$, S.D.=7.8 Hz; frequency range $\bar{x}=41.7$, S.D.=14.9) of relatively short duration ($\bar{x}=1.2$, S.D.=0.3 s). Approximately half of the signals exhibited a characteristic s-shape pattern when examined spectrographically. Long duration sweeps and pulses, typical of blue whales sounds from the eastern North and Central Pacific, were not detected. The vocalizations recorded in this study are compared to recordings of short duration blue whale sounds made elsewhere, including off California, Baja California, and southwestern Australia. In general, short duration blue whale sounds from these other areas are more variable and/or have different frequency characteristics. The function of short duration, FM blue whale sounds is unknown.

8:45

5aAB3. Sterotyped and other vocalizations of fin whales (*Balaenoptera physalus*) off Chile. Paul O. Thompson (3845 Falcon St., San Diego, CA 92103) and William C. Cummings (Oceanogr. Consultants, San Diego, CA 92122)

Underwater vocalizations from a small group of fin whales were recorded from the U.S. National Science Foundation's (NSF) research vessel ''HERO'' in Golfo de Penas, Chile, May, 1970. The most outstanding and frequent vocalizations were patterned utterances of about 1.5-s duration that swept upward from about 35 to 65 Hz. Endings typically were emphasized with an outstanding increase in sound level. The second most frequent type of vocalization was 1-s downward sweeps in a range of about 100–30 Hz. Both types occurred in small groupings with evenly spaced intersound intervals of about 8.5 s (upsweeps) and about 4 s (downsweeps). These sterotyped vocalizations had acoustical characteristics that were very similar to those of sounds that we had recorded from fin whales in the Gulf of California (Mexico) during 1969. [Work supported by NSF.]

9:00

5aAB4. Deep, harmonic moans associated with Bryde's whales in several locations worldwide. Jay Barlow (Southwest Fisheries Sci. Ctr., P.O. Box 271, La Jolla, CA 92038), Erin Oleson (Scripps Inst. of Oceanogr., La Jolla, CA), and Mark McDonald (2535 Sky View Ln., Laramie, WY 82070)

Several sounds have been associated with Bryde's whales in the Gulf of California [Cummings *et al.* (1986); Edds *et al.* (1993)]. Here, we report on harmonically rich tonal calls recorded in the vicinity of Bryde's whales. In the eastern tropical Pacific, sonobuoys were deployed in the immediate vicinity of Bryde's whales on 40+ occasions, and in two cases, a 1–2-s tonal signal was recorded with a fundamental of 50–53 Hz and with 3 clear harmonic overtones. In the southern Caribbean, sonobuoys were deployed in the vicinity of Bryde's whales, and, on 20 occasions, a 2–3-s tonal signal was recorded with a fundamental of either 15–20 or 40–45 Hz and with 3–5 clear harmonic overtones. Off New Zealand, a bottom-mounted hydrophone recorded 4–7-s, 22–28-Hz tonal calls which are seasonally associated with the presence of Bryde's whales; such calls were previously recorded in this area using a hydrophone with greater

bandwidth [Kibblewhite *et al.* (1967)] and showed harmonic overtones that were similar to those at the other two locations. Bryde's whales appear to produce a rich variety of calls, including these harmonic moans that had not been previously noted in the published literature.

9:15

5aAB5. Acoustic identification of nine delphinid species in the eastern tropical Pacific Ocean. Julie N. Oswald, Jay Barlow (Southwest Fisheries Sci. Ctr., Natl. Marine Fisheries Service, NOAA, P.O. Box 271, La Jolla, CA 92038-0271, joswald@mail.sdsu.edu), and Tom Norris (SAIC, San Diego, CA 92110)

Acoustic techniques have the potential to increase the reliability of cetacean species identification during shipboard surveys. The whistles of nine odontocete species were compared using data collected from a towed array and sonobuoys deployed during dolphin abundance surveys in the eastern tropical Pacific. Twelve variables were measured manually from spectrographic displays of each whistle ($n=912$). Multivariate discriminant function analysis (DFA) resulted in 49.9% of whistles being classified to the correct species. It was hypothesized that some whistles carry less species-specific information than others, therefore, groups of five whistles were averaged to reduce the effect of these ambiguous whistles. Correct classification increased to 65.4% when DFA was run on the averaged data set. A species identification decision tree that used 7 of the 12 whistle variables was constructed using nonparametric techniques (classification and regression trees) and resulted in 53.1% correct classification when applied to the original data set. Prior probabilities were added to the decision tree based on sighting rates for each species in the study area, resulting in 56.7% correct classification. The species identification decision tree provides a relatively simple acoustic method that can be used to augment conventional visual techniques.

9:30

5aAB6. Geographic variations in the whistle repertoire of Hawaiian spinner dolphins (*Stenella longirostris*). Carmen Bazua-Duran and Whitlow Au (MMRP, Univ. of Hawaii, P.O. Box 1106, Kailua, HI 96734, bazua@soest.hawaii.edu)

Studying geographic variations in the whistle repertoire of Hawaiian spinner dolphins (*Stenella longirostris*) will increase our understanding of the population structure of this species. In the present work, groups of spinner dolphins off the Islands of Oahu, Lanai, and Hawaii, were studied. The dolphin whistle repertoire of each island was used to search for geographic variations. Two methods were used: (1) spectrogram information extraction and (2) zero crossing. For the first method, frequency and time information was extracted from the spectrogram of each whistle. Results show that geographic variations exist in the whistle repertoire of Hawaiian spinner dolphins. Statistically significant differences were obtained by using discriminant function analysis: Wilks lambda=0.576; $F(22,1866)=26.93$, $p<0.00001$, N(Oahu)=314, N(Lanai)=293, and N(Hawaii)=339. Greater differences were found between the Oahu and Hawaii groups (3.413), intermediate differences were found between the Oahu and Lanai groups (2.200), and Lanai and Hawaii were the least different groups (0.560). These results indicate that looking for geographic variations in the whistle repertoire will help us in understanding the population structure of Hawaiian spinner dolphins. The results of the zero-crossing method will be compared to the spectrogram information extraction results. [Work supported by Leonida Memorial Scholarship, UH Seed Money Grant, UH Foundation Grant, Fulbright-GarcaRobles-CONACyT and DGAPA.]

9:45

5aAB7. Acoustic analysis of objects ensonified by a bottlenose dolphin during a cross-modal matching task. Caroline M. DeLong, Whitlow W. L. Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734, delong@hawaii.edu), and Heidi E. Harley (New College of the Univ. of South Florida, Sarasota, FL 34243)

A bottlenose dolphin performed a matching task in which he was presented with an object (e.g., tin cup) in one modality (e.g., vision) and then asked to choose the same object from a group of three objects using another modality (e.g., echolocation) or the same modality (e.g., vision). Object sets were presented in two cross-modal conditions (visual sample with echoic choices or echoic sample with visual choices) and two intramodal conditions (visual sample with visual choices or echoic sample with echoic choices). These objects were later ensonified by dolphin-like clicks from several different angles and the signals were captured and digitized (objects were aspect dependent, thus producing different echoes at different angles). Echoes from object sets on which the dolphin performed well were compared with echoes from objects sets on which the dolphin performed poorly to elucidate the aspects of the echo signals that the dolphin used to discriminate among the objects (e.g., echo highlights, spectral characteristics) and the reason(s) why some objects were easier to tell apart than others. Visual analysis of the echolocation behavior and search strategy of the dolphin using video recordings supplemented the acoustic analysis.

10:00–10:15 Break

10:15

5aAB8. Echolocation in the Risso's dolphin, *Grampus griseus*. Jennifer D. Philips, Paul E. Nachtigall, Whitlow W. L. Au, Jeffrey L. Pawloski (Marine Mammal Res. Prog., Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734, jphilips@hawaii.edu), and Herbert L. Roitblat (Univ. of Hawaii at Manoa, Honolulu, HI 96822)

This study investigated whether the Rissos dolphin, *Grampus griseus*, echolocates. The Rissos dolphin is a gregarious, near-shore pelagic delphinid that consumes nearly exclusively cephalopods and has a distinctive groove along its forehead. In this study, an individual female Rissos dolphin was trained to discriminate an aluminum cylinder from a nylon sphere (Experiment 1) or an aluminum sphere (Experiment 2). The animal wore eyecups and was required to swim to and touch the standard cylinder. The animals choice accuracy was 98.3% in Experiment 1 and 99.6% in Experiment 2, indicating that she was able to perform the target discriminations with little difficulty despite being blindfolded. Clicks were acquired at amplitudes up to 202 dB ($M=199$ dB), peak frequencies up to 105 kHz ($M=51$ kHz), center frequencies up to 70–75 kHz ($M=57$ kHz), bandwidths of 31 to 92 kHz, and durations of 40–70 ms. The data establish that the Rissos dolphin echolocates and that its clicks are very similar to the clicks of other echolocating dolphins. Implications of the unique behavioral and physical characteristics of this species on its sonar system are discussed.

10:30

5aAB9. Extraction and tracking of bottlenose dolphin whistle contours. Ryuji Suzuki and John R. Buck (Dept. of ECE & CMaST, UMass Dartmouth, 285 Old Westport Rd., N. Dartmouth, MA 02747)

The time-frequency contours of bottlenose dolphin (*Tursiops truncatus*) whistles are commonly extracted as features for the classification and comparison of these signals. This research presents a new method to extract and track the whistles' fundamental frequency contours. This method takes a spectrogram as an input and uses a Bayesian inference approach with a random walk model for the frequency transitions. The output of the method is the maximum *a posteriori* probability estimate of the frequency contour. This technique is similar to the Kalman filter except that the states are probability densities, and it uses a probabilistic state transition model. For each time segment, the method combines the prior estimate of the probability distribution with the current observation to obtain the current estimate of the probability distribution. Experimental evaluation indicates that this method greatly improves the reliability of contour extraction from noisy signals when compared to the *ad hoc* contour extractor proposed by Buck and Tyack [J. Acoust. Soc. Am. **94** (1993)]. This technique has previously applied to target tracking in radar [Bethel and Rahikka, IEEE Trans. Aerosp. **23**(6) (1987)]. [Work supported by NSF Ocean Sciences.]

10:45

5aAB10. Autoregressive synthesis of bottlenose dolphin whistles. Xiaozhou Huang and John R. Buck (Dept. of ECE & CMaST, UMass Dartmouth, 285 Old Westport Rd., N. Dartmouth, MA 02747)

The signature whistle hypothesis proposes that most of the whistles bottlenose dolphins (*Tursiops truncatus*) produce are individually distinctive, and that bottlenose dolphins can identify the signature whistles of the other dolphins from the fundamental time-frequency contour. This hypothesis has not been tested directly. One major obstacle has been the need to modify the whistles so that scientists can quantitatively control the acoustic parameters of the whistles. Buck, Morgenbesser, and Tyack [J. Acoust. Soc. Am. **108** (2000)] proposed a model for synthesizing whistles using a sum of harmonically related sinusoids. Some synthetic whistles produced by this method sound artificial. The objective of the current research is to develop an alternative technique for modifying and synthesizing these whistles using autoregressive (AR) modeling. An important property of the AR model is that if the synthesis filter is excited by the residual signal from a prediction filter, the original signal is perfectly reconstructed. Using the residual signal as the excitation for a synthesis filter with modified AR parameters should produce more realistic-sounding modified whistles. [Work supported by NSF Ocean Sciences.]

11:00

5aAB11. Horizontal angular discrimination by an echolocating bottlenose dolphin (*Tursiops truncatus*). Brian K. Branstetter, Louis M. Herman, Adam A. Pack, Sonja J. Mevissen, Amanda M. Moore, Jamie Granum, Lea Carsrud, Carol Butler, Sarah B. Lecaroz, and Scotty Roberts (Kewalo Basin Marine Mammal Lab., Univ. of Hawaii, Manoa, 1129 Ala Moana Blvd., Honolulu, HI 96814, branstet@hawaii.edu)

A bottlenose dolphin was tested on its ability to echoically discriminate horizontal angular differences between two arrays of vertical, air-filled PVC rods. The blindfolded dolphin was required to station in a submerged vertically oriented hoop 2 radial meters from the stimuli and indicate if an array with four rods ($S+$) was to the left or the right of an array with two rods ($S-$) by pressing a corresponding paddle. The rods within each array were separated by 2 deg and the two arrays were separated by eight different angles between 2.25 and 6 degrees. A modified method of constant stimuli was used to test for angular discrimination ability. The results yielded a high-pass psychometric function with an arbitrary 75% correct threshold of 1.5 deg. These data agree well with passive listening minimum audible angle thresholds of 0.9 deg for click signals and 2.1 deg for a pure tone signal (Renaud and Popper, 1975). Analyses of response times, number of clicks and inter-click intervals suggested no significant adaptive behavior was used as the task became more difficult. These results help define angular resolution capabilities of dolphin sonar that may play an important role in representing spatial information in the dolphin's environment.

11:15

5aAB12. Hector's dolphin (*Cephalorhynchus hectori*) vocalizations and gillnet pingers. Laura Cavagnaro, Kenneth Baldwin (Chase Ocean Eng. Lab., 24 Colovos Rd., Durham, NH 03857), and Gregory Stone (New England Aquarium, Central Wharf, Boston, MA 02110)

The effectiveness of acoustic gillnet pingers for eliciting a deterrent behavior in some marine mammals is well documented, although the reasons for this behavior are not clearly understood. Little is known about the mammals vocal response to these devices. The objective of the experiments defined here was to investigate the possibility of vocalization response changes due to the presence of a gillnet pinger. Three commercially available gillnet pingers were used as acoustic cues for Hectors dolphins (Cephalorhynchus hectori). The dolphins' vocalizations were recorded with and without the pinger in the water. The records of the dolphin vocalizations were analyzed using digital signal processing techniques to determine whether the dolphin vocalizations changed with the presence of an acoustic cue from the pinger. Fourier analysis, wavelet analysis, and empirical orthonormal functions were used to analyze the vocalizations. No significant changes in the vocalizations could be found between the control and active gillnet pinger cases.

FRIDAY MORNING, 8 DECEMBER 2000 PACIFIC SALON A, 8:25 TO 11:45 A.M.

Session 5aEA

Engineering Acoustics: Special Purpose Acoustic Sources

Steven R. Baker, Chair

Department of Physics, Naval Postgraduate School, Code PH/Ba, Monterey, California 93943

Chair's Introduction—8:25

Contributed Papers

8:30

5aEA1. Broadband amplitude enhancement of a transducer using parametric excitation. Jerry H. Ginsberg and Phillip H. Nguyen (G. W. Woodruff School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405)

A paper by Larson *et al.* [J. Acoust. Soc. Am. **103**, 1428–1441 (1998)] suggests that it is possible to use piezoceramics to construct a transducer with a time-variable spring constant. The present work explores the response of such a transducer, modeled as a one-degree-of-freedom system, when the stiffness is varied with a fixed period either sinusoidally or as a square wave. Such variation represents a parametric excitation whose effect strongly modifies the conventional forced response. Transient solutions of the equations of motion are obtained by numerical integration, and verified in the square wave case by an analytical solution that exploits the piecewise linear nature of the equations. It is shown that the response amplitude is increased by as much as 20 dB over a range of excitation frequencies from DC to twice the nominal resonant frequency when the period of the parametric excitation is set substantially below the natural frequency.

8:45

5aEA2. Use of actively tuned low-frequency transducers for high-power linear chirp transmission. Julien Bernard and George A. Lesieutre (Dept. of Aerosp. Eng., Penn State Univ., 233 Hammond Bldg., University Park, PA 16802)

To overcome the low-radiation resistance, low-frequency underwater transducers must be operated in the vicinity of resonance. To provide maximum displacement at resonance, they are often designed as high Q, hence narrow bandwidth, devices. Various approaches have been used to

increase the bandwidth of these transducers, including impedance matching layers and velocity feedback, among others. However these approaches result in reduced peak sensitivity. In contrast, low-frequency transducers with a variable and controllable resonance frequency, which may be matched to the transmitted signal dominant frequency at all times, offer the double advantage of increased bandwidth and high sensitivity at almost all frequencies within the bandwidth. Variable frequency transducers have been used for the high-power transmission of consecutive single frequency pulses at various discrete frequencies, such as stepped frequency modulated chirps. This paper shows that they can also provide high-power transmission of continuously varying frequency pulses, such as linear frequency modulated chirps. Using a simple equivalent circuit model, the responses of a fixed frequency and a variable frequency piezoceramic tonpilz to a linear chirp are calculated and compared, in the time and the frequency domains. The results show that the variable frequency transducer can efficiently transmit linear chirps of much larger bandwidth.

9:00

5aEA3. Design of actively tuned flexural piezoceramic bar/disk transducers. Julien Bernard and George A. Lesieutre (Dept. of Aerosp. Eng., Pennsylvania State Univ., 233 Hammond Bldg., University Park, PA 16802, jeb21@kirkof.psu.edu)

Underwater transducers with variable and controllable resonance frequency can efficiently transmit stepped or linear chirp signals of greater bandwidth by continuously matching their resonance frequency to the dominant frequency of the transmitted signal. This paper presents a technique to vary the first resonance frequency of some widely used underwater transducers: flexural piezoceramic bars and disks. Positive-bias electric fields are added to the driving ac field to generate either a tensile in-plane load in k_{31} coupling, or a compressive in-plane load in k_{33} coupling. This load modifies the flexural rigidity of the transducer, which in turn affects its resonance frequencies. Ideally, supported transducers are first analyzed, from the point of view of the electrochemical coupling coefficient at the first resonance, and the frequency shift per applied bias field. Transducers using elastic hinges to approximate simply supported boundary conditions are then examined. The response of the transducers to the combined dynamic bias field and driving ac field is simulated to look for possible dynamic effects, such as the excitation of high-frequency extension and bending modes. Finally, experimental measurements of the coupling coefficient and frequency shift per bias field of prototype transducers are presented. [Work supported by ONR.]

9:15

5aEA4. Philip Reis, telephone pioneer. Robert T. Beyer (Dept. of Phys., Brown Univ., Providence, RI 02912)

The name of Phillip Reis is widely known and respected in Germany for his early work on the telephone, but he is almost unknown in America. A teacher at the high school level, Reis put together both a mouthpiece and a receiver, and called the combination a telephone, demonstrating its operation on a number of occasions. This was in 1863, 13 years before Bell's invention. While the U.S. Supreme Court ruled in 1887 that ''to follow Reis is to fail, to follow Bell is to succeed,'' it must be remembered that Reis died in 1873, and therefore, had no opportunity to rebut the arguments of Bell and his associates. The distinguished Sylvanus Thompson, Professor of Physics at the University of Bristol, wrote of Reis in 1883, ''The inventor of The Telephone (Reis) will be remembered and honoured in the coming if not in the present age.'' A discussion of Reis's instrument, including various comments upon it, will be given.

9:30

5aEA5. Beam pattern formation and on-/off-axis pressure in the near field through elastic structures. Fernando García-Osuna, Elmer L. Hixson (Dept. of Elec. and Computer Eng., Univ. of Texas, Austin, TX 78712), and Augusto L. Podio (Univ. of Texas, Austin, TX 78712)

Detailed measurements have been done and compared to analytical models of the beam pattern and on-/off-axis pressure in the near-field region of plane circular transducers in order to model the reflection and transmission of ultrasonic waves from and through steel pipes (casing) and plates immersed in oil. It is shown that for broadband and pure tone pulses in the near field, for free field conditions and through a curved or planar layered media, the sidelobes are smaller and the half-power beamwidth is narrower than in the far field. The beamwidth is slightly dispersed (<1 deg) through the casing and plates of similar thickness, being larger through the planar geometry. Off-axis pressure through these elastic structures reveals condensing and dispersing effects (<2 mm) in the curved and planar geometry, respectively. On-axis pressure, however, for $\lambda <$ structure's thickness and plane ultrasonic transducers with size $a/\lambda > 5$, where a is the transducer radius, the reflected and transmitted sound field is the same from and through both structures and behaves as a well-collimated beam in the near-field region. [Work supported by the University of Texas at Austin and the Mexican Petroleum Institute.]

9:45

5aEA6. Broadband barrel-stave flextensional transducers. Dennis F. Jones (Defence Res. Establishment Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada)

Class III barrel-stave flextensional transducers are compact underwater sound sources that have broadband capabilities arising from close coupling between flexural and longitudinal vibrational modes. The electroacoustic performances of several of these transducers were measured at both the Defence Research Establishment Atlantic Acoustic Calibration Barge on Bedford Basin and the Seneca Lake Sonar Test Facility operated by the Naval Undersea Warfare Center Division Newport. Transmitting voltage response curves were obtained over the frequency range 0.5 to 20 kHz, and vertical directivity patterns at selected frequencies were plotted. A stress-limited source level of 204 dB//1μPa-m at 5 kHz was determined from high-power tests. Since the transducers were calibrated at water depths ranging from 15 to 107 m, the effects of hydrostatic pressure on the resonance frequencies were quantified. Recently, the Class III transducer was integrated into a battery-operated Broadband Acoustic Transmission System (BATS) which is currently being used for broadband sonar and marine mammal playback studies. A brief overview of the BATS hardware is presented. [Work supported in part by ONR.]

10:00–10:15 Break

10:15

5aEA7. Analysis and optimal design and a class flextensional transducer. Yongrae Roh and Kukjin Kang (Dept. of Sensor Eng., Kyungpook Natl. Univ., Sankyug-Dong 1370, Taegu 702-701, Korea, yryong@eeg.kyungpook.ac)

With the finite-element method, we analyzed variation of the sound pressure level and internal thermal distribution of a class flextensional transducer in relation to various design factors in its structures and material properties. The design factors included lengths of major and minor axes, thickness and height of a shell, length and width of a PZT stack, shape of an edge insert, thickness of a rubber coating, and all their material properties. Effects of each of the factors on the transducer performance were analyzed in a quantitative manner as a function of frequency. Based on the results, we determined the optimal structure of a class flextensional transducer that had 1-kHz resonance frequency, the maximum sound pressure level, and the minimum temperature rise in operation. The sound pressure by the optimal structure was higher than that of a basic structure by two times, and the temperature level in operation was lower by 20%. Results of the present work can be utilized to design class flextensional transducers of various resonance frequencies.

10:30

5aEA8. A modal Pritchard approximation for computing array element mutual impedance. Clyde L. Scandrett (Dept. of Mathematics, Naval Postgrad. School, Monterey, CA 93943), Joe L. Day (Office of Naval Intelligence), and Steven R. Baker (Naval Postgrad. School)

An investigation into the applicability and accuracy of Pritchard's approximation for closely packed transducer arrays is undertaken. A new, ''modal'' Pritchard approximation is developed, based upon normal modes of the acoustic medium, and is tested for arrays of acoustically hard

spheres to ascertain its accuracy in determining the mutual acoustic radiation impedance between array elements. For $ka \approx 1$, it is found that the modal Pritchard approximation works quite well in approximating the mutual radiation impedance of a two-element array, even for relatively close spacing, but for arrays of three or more scatterers in close proximity, the approximation may have relatively large errors. The effect of neglecting interelement scattering is analyzed for the monopole-to-monopole scattering of various configurations of a three-element array and a 16-element double-line array. [Work supported by ONR.]

10:45

5aEA9. Directivity of conformal-baffled cylindrical transducers. John Welch, Lawrence Reinhart, Tetsuro Oishi (Acoust. Res. Lab., Elec. Eng. and Adv. Technol. & Manuf. Ctr., Univ. of Massachusetts, Dartmouth, N. Dartmouth, MA 02747), Boris Aronov, and David A. Brown (BTECH, East Providence, RI 02915)

The theoretical and experimental directivity patterns, radiation impedance, and source levels of piezoelectric cylindrical rings with conformal acoustic baffles will be presented. A model was implemented with MATLAB to simulate the directivity response of various baffled cylindrical piezoelectric transducers. The cylinder is modeled by assuming a vibrating cylindrical element with constant surface velocity as outlined by Morse [*Vibration and Sound*, 3rd ed.] and adding complications due to effects of the conformal acoustic baffle. A number of devices of practical importance have been fabricated and tested. Results for conformal baffles ranging from 180 to 250 deg of coverage are presented and are compared with experimental results using different baffle designs and at various frequencies.

11:00

5aEA10. Active underwater acoustic cancellation with pressure sensor and piezocomposite transducer. Jinmin Zhang (DSO National Labs., 20 Science Park Dr., Singapore 118230), Woosuk Zhang, Vijay K. Varadan, and Vasundara V. Varadan (Ctr. for the Eng. of Electron. and Acoust. Mater., Penn State Univ., State College, PA 16802)

Active underwater acoustic cancellation is investigated experimentally in a water-filled pulse tube using a pressure sensor and a piezo-composite actuator. The pressure sensor is made of thin 1-3 piezo-composite disk encapsulated in acoustic transparent rubber and put right in front of the actuator. A 20-dB echo reduction to harmonic incident waves in 5–10 kHz is achieved experimentally by directly phase-shifting the sensor signal and amplifying it to drive the canceling actuator. A robust analog controller is also developed based on the experimentally identified acoustic reflection coefficient of the piezo-composite transducer with steel backing and the transfer function from piezo-composite transducer to the pressure sensor. Issues such as robustness and stability of the analog controller are also discussed.

11:15

5aEA11. Nonlinear propagation of multitone one-dimensional waves. Application to sound systems modeling and analysis. Alexander Voishvillo (Cerwin-Vega, Inc., 555 E. Easy St., Simi Valley, CA 93065, avoishvillo@iccas.com)

Multitone stimulus is used in identification of weakly nonlinear systems and for assessment of nonlinearity in sound and broadcasting equipment. Multitones' spectral and statistical characteristics resemble those of the musical signals; the nonlinear response to multitone stimulus yields more versatile information about nonlinearity than conventional measurements of harmonic and two-tone intermodulation distortion. Earlier work has demonstrated that strong harmonic and intermodulation distortions accompany the propagation of waves radiated by large horn arrays. In this current work, the models of propagation of multitone waves are developed. The modeling considers propagation of multitone one-dimensional waves. The modeling is based on numerical solution of implicit equations describing distorted one-dimensional waveforms. A nonlinear model of multitone waves propagating within a horn is developed. The model considers linear propagation of the fundamental tones with reflections from the mouth and the nonlinear, reflection-free propagation of distortion products. Further research may include radiation of multitone waves by a complex source and substitution of multitone stimulus by a musical signal to provide listening tests and auralization of propagation nonlinearity. The data obtained from modeling may be used to build Volterra models of nonlinear propagation. These models might lead to signal processing solutions to minimize distortion in propagating waves.

11:30

5aEA12. Hysteresis effects in high-power ultrasonic transducers. Bruce J. P. Mortimer, Alex Green, Tertius du Bruin, Ian de Vries (Ctr. for Instrumentation Res., Cape Technikon, P.O. Box 652, Cape Town 8000, South Africa), and Jon T. Tapson (Univ. of Cape Town, Rondebosch 7700, South Africa)

It is well known that ultrasonic piezoelectric transducers can be difficult to drive at high power levels. Transducers usually have high Qs, which implies a narrow driving bandwidth and characteristics that may change with temperature, component wear, or variations in the load. Some method of feedback control is usually required to lock onto the optimum transducer driving frequency. Small-signal parameters are usually used in the design of both the control and the amplifier matching system. In this work we have measured the characteristics of several sandwich transducers coupled to water at significant drive levels (50–250 V) by measuring the drive current, voltage, and phase. Increasing the drive level shows a decrease in optimum operating-point frequency. In addition, both the Q and admittance at the driving point decreases. This creates a practical problem as the matching of transducers done on the basis of small-signal values will be inaccurate. Increasing the power levels also results in an abrupt transition to a lower impedance value after the onset of cavitation. Decreasing the power levels is shown to exhibit a hysteresis effect which is thought to be due to the persistence of gas bubbles in the liquid.

Session 5aPA

Physical Acoustics: Scattering—Periodic and Aperiodic

John A. Burkhardt, Chair
Mechanical Engineering, U.S. Naval Academy, 590 Holloway Road, Annapolis, Maryland 21402

Contributed Papers

8:15

5aPA1. Localization of acoustic waves in 1-D random liquid media. Pi-Gang Luan and Zhen Ye (Wave Phenomena Lab., Dept. of Phys., Natl. Central Univ., Chungli, Taiwan 320, ROC, luan@fermi.phy.ncu.edu.tw)

We study acoustic propagation in one-dimensional water ducts containing many air-filled blocks. The acoustic band structures for the periodic arrangements of the blocks are calculated, whereas the transmission for various random configurations of the blocks is computed by the transfer matrix method. It is shown that waves at all frequencies become localized even with a small amount of randomness. The spatial distribution of the localized energy is investigated, and is shown not to be trapped near the source, contrary to higher dimensional cases. The results also reveal a distinct collective behavior for localized waves, a feature useful for distinguishing the localization from the residual absorption effect.

8:30

5aPA2. Band gaps and localization in periodic and random arrays of 2-D isotropic acoustic scatterers. Emile M. Hoskinson (Dept. of Phys., Univ. of California, Berkeley, Berkeley, CA 94720-7300) and Zhen Ye (Nat. Central Univ., Chungli, Taiwan 320, ROC, hoskins@socrates.berkeley.edu)

A band-structure calculation for a periodic array of scatterers reveals band gaps—bands of frequency outside of which propagation is relatively unhindered, inside of which waves are evanescent and cannot propagate away from their source. In the case of an array of resonant isotropic 2-D acoustic scatterers, a direct, exact, self-consistent technique incorporating all orders of multiple scattering can be used to calculate the scattered field due to a given source. This can be done for the periodic case where the band-structure calculation is relevant: a region of frequencies is found in which the scattered field is localized around the source, and this region coincides with the calculated band gap. The self-consistent technique has the advantage that it can also be used for an arbitrary random distribution of scatterers. Surprisingly, the localization region proves relatively independent of the scatterer distribution. Further investigation reveals that as the frequency is brought within the localization region, the scatter oscillation phases undergo a kind of phase transition into a coherent state which persists within the localization region, also appearing regardless of the scatterer distribution. [Work supported by the National Science Council of the ROC and National Central University.]

8:45

5aPA3. Acoustic scattering by arrays of air bubbles in water. Zhen Ye (Wave Phenomena Lab., Dept. of Phys., Natl. Central Univ., Chungli, Taiwan 320, ROC)

Wave localization is a ubiquitous phenomenon. It refers to situations that transmitted waves in scattering media are trapped in space and remain confined in the vicinity of the initial site until dissipated. Here we report a phase transition from acoustically extended to localized states in arrays of identical air-filled bubbles in water. It is shown that the acoustic localization in such media is coincident with the complete band gap of a lattice arrangement of the air bubbles. When the localization or the band gap occurs, a peculiar collective behavior of the bubbles appears. The acoustic transmission through a planar array of air bubbles is also studied.

9:00

5aPA4. Acoustic stop bands for two-dimensional periodic arrays of metallic rods: Honeycomb geometry. M. S. Kushwaha (Inst. of Phys., Univ. of Puebla, P.O. Box J-45, Puebla, Pue. 72570, Mexico, manvir@sirio.ifuap.buap.mx) and B. Djafari-Rouhani (Univ. of Sci. & Technol. of Lille-I, 59655 Villeneuve D'Ascq Cedex, France)

Extensive band structures have been computed for periodic arrays of rigid metallic rods in air for honeycomb structure. Multiple complete acoustic stop bands have been obtained within which sound and vibrations are forbidden. These gaps start opening up for a filling fraction $f \geqslant 8\%$ and tend to increase with the filling fraction, exhibiting a maximum at the close-packing. A tandem structure has also been proposed that allows an ultrawideband filter for environmental or industrial noise to be achieved in the desired frequency range. This work is motivated by the recent experimental observation of sound attenuation on the sculpture by Eusebio Sempere, exhibited at the Juan March Foundation in Madrid [Nature **378**, 241 (1995)] and complements the corresponding theoretical work (see the preceding abstract). [Work partially supported by CONACyT Grant No. 28110E.]

9:15–9:30 Break

9:30

5aPA5. Acoustic stop bands for two-dimensional periodic arrays of metallic rods: Square geometry. M. S. Kushwaha (Inst. of Phys., Univ. of Puebla, P.O. Box J-45, Puebla 72570, Mexico, manvir@sirio.ifuap.buap.mx)

Extensive band structures have been performed for two-dimensional periodic arrays of rigid stainless steel cylinders in air, with Bloch vector being perpendicular to the cylinders. For cylinders 2.9 cm in diameter and period of 10 cm (i.e., the filling fraction $f = 0.066$) there is no acoustic gap for frequencies below 6.4 kHz. However, the density of states reveals prominent minima at 1.7 and 2.4 kHz. These frequencies do agree with those of the first two attenuation maxima observed in the experiment [Nature **378**, 241 (1995)] and are indeed related to diffraction from the [100] and [110] planes. Thus even with idealization, Sempere's sculpture exhibits pseudogaps—not full gaps. It is stressed that, for any value of the period of the system, there is no acoustic gap for $f \leqslant 30\%$; magnitude of the gap, for $f \geqslant 30\%$, is inversely proportional to the period of the system. Moreover, if the filling fraction exceeds 40%, a second gap, higher in frequency, opens up. In addition, the fabrication of a multiperiodic system in tandem that could create a huge hole in sound within the human audible range of frequencies is proposed. [Work partially supported by CONCyT Grant No. 28110E.]

9:45

5aPA6. Sound control by two-dimensional periodic structures. J. Sánchez-Dehesa, D. Caballero (Dept. of Theoretical Condensed Matter, Autonomous Univ., 28049 Madrid, Spain), R. Martínez-Sala, C. Rubio, J. V. Sánchez-Pérez, L. Sanchis, and F. Meseguer (Polytech. Univ., 46022 Valencia, Spain)

Sound transmission properties of two-dimensional triangular arrays of rigid cylinders in which we created a lattice of vacancies with rectangular symmetry will be presented. At low frequencies, the attenuation spectra measured along the high-symmetry directions show an attenuation band corresponding to the lattice of vacancies. On the other hand, at high frequencies, the attenuation bands associated with the full triangular lattice still remain. This result indicates that the vacancies act as a true lattice of scatterers embedded in a triangular lattice. Acoustic band structures have been calculated by a variational method and compared with attenuation spectra. The features of the spectra are well described by the band calculation. An intuitive approach that explains the experimental results will be given. [Work supported by CICyT of Spain.]

10:00

5aPA7. Sonic band-gap materials as acoustical components. F. Cervera, L. Sanchis, R. Martínez-Sala, C. Rubio, J. V. Sánchez-Pérez, F. Meseguer, C. López (Unidad Asociada CSIC-UPV, Dept. of Appl. Phys., Polytech. Univ., 46022 Valencia, Spain), D. Caballero, and J. Sánchez-Dehesa (Autonomous Univ., 28049 Madrid, Spain)

Sonic band-gap (SBG) crystals can be made by periodic distributions of scatterers. In particular, it was found that two-dimensional (2D) arrays of rigid cylinders in air do have acoustic band gaps that could be of application in acoustic devices related to control and filtering of sound. Here, it is pointed out that this type of SBG crystal can be used as custom-made acoustical components in the linear region well below the gap. This idea is supported by the fact that sound velocities in these materials are smaller than that of air. Also, it was previously measured that their transmittance is high; i.e., they exhibit low reflectance [J. V. Sánchez-Pérez *et al.*, Phys. Rev. Lett. **80**, 1208–1212 (1998)]. Both features guarantee the practicality of sonic crystal lenses and other optical components. Finally, 2D SBG crystals based on triangular and square arrays are classified as acoustically uniaxial or biaxial. [Work supported by CICyT of Spain.]

10:15

5aPA8. Rough surfaces modeling using CTFM sonar. Konstantinos Zografos and Penny Probert (Dept. of Eng. Sci., Oxford Univ., Parks Rd., Oxford OX1 3PJ, UK)

A new statistical model for the scattering of acoustic signals from rough surfaces is presented. It has been developed to support texture measurements from walls and other structural surfaces using air-borne sonar. The rough surface is divided into a number of elementary surfaces, each with random orientation and positioning around the mean for the whole surface. Each elementary surface is modeled using an extension of the Kirchhoff approximation method for a monochromatic wave source [see O. Bozma and R. Kuc, J Acoust. Soc. Am. **89**, 2519–2531 (1991)]. Whereas previous work has accounted only for the envelope of the overall energy received at different angles, from a rough surface, the new model accounts for the nonuniformity in the energy pattern. The model is applied to describe the signal received by a KASPA frequency-modulated sensor, a sensor developed originally as a mobility aid for the blind [see L. Kay, *Electronic Spatial Sensing for the Blind* (Martinus Nijhoff, Boston, 1985), pp. 125–139]. It has been validated with data from different textured surfaces on buildings in Oxford. In the paper we will discuss its use in classifying surfaces to assist navigation and for inspection.

10:30–10:45 Break

10:45

5aPA9. Characterization of multiple scattering acoustic media using backscattered pulsed ultrasound. John Burkhardt (U.S. Naval Acad., 11C Annapolis, MD 21402)

A technique is presented for the determination of the Legendre moments of the differential scattering cross section per unit volume of a multiple scattering acoustic medium. Acoustic propagation within the medium is modeled using a time-dependent diffusion approximation to the scalar radiative transfer equation. An expression for the backscattered acoustic intensity is derived. Determination of the Legendre moments requires backscatter intensity measurements as a function of azimuthal angle at a fixed polar angle. Monte Carlo simulations are presented which support the proposed technique.

11:00

5aPA10. Generalized optical theorem for acoustic backscattering. Philip L. Marston (Dept. of Phys., Washington State Univ., Pullman, WA 99164-2814, marston@wsu.edu)

The usual optical theorem for sound expresses the extinction cross section as proportional to a component of the far-field forward scattering amplitude for three-dimensional scattering [V. Varatharajulu, J. Math. Phys. **18**, 537–543 (1977); S. G. Kargl and P. L. Marston, J. Acoust. Soc. Am. **88**, 1103–1113 (1990)]. It does not appear to be widely recognized, however, that in the absence of dissipation, a component of the acoustic scattering at other scattering angles may be related to an angular integration involving scattering amplitudes in other directions (related by supplemental conditions to the direction of interest) provided the target has inversion symmetry. This analytical result is confirmed computationally for the imaginary component of the form function for backscattering by spheres described by a partial-wave series where the normalization of Kargl and Marston is used as well as numerical integration. The usual optical theorem is the analytical result for the forward scattering case. [Work supported by the Office of Naval Research.]

11:15

5aPA11. Finite scatterer volume corrections to multiple scattering theories. Douglas Photiadis (Naval Res. Lab., 4555 Overlook Ave. SW, Washington, DC 20375-5000)

Multiple scattering theories used to describe random acoustic media usually assume the scatterers within the medium are pointlike and statistically independent from one another. The finite volume of the scatterers may, however, be relevant when the density of the scatterers is large. A general theoretical approach in the context of multiple scattering theory has been developed to account for this effect without introducing correlations between the scatterers. Employing this scheme, the effective wave speed, correlation length, and density of states have been evaluated for both two- and three-dimensional media and compared to point scatterer results. The corrections are found to be of the order of the fraction of the total volume occupied by the scatterers, and hence relevant to attempts to extend multiple scattering theories into the high density regime. Analytic results are obtained in both bulk acoustic and structural acoustic systems. [Research supported by ONR.]

11:30

5aPA12. A comparison between one-channel time reversal and phase conjugation in random media and in reverberant cavities. Mathias Fink, Arnaud Tourin, Julien de Rosny, and Arnaud Derode (Laboratoire Ondes et Acoustique, ESPCI, Univ. of Paris, 7 10 rue Vauquelin, 75005 Paris, France)

In this paper, we are interested in time-reversal experiments conducted with only one omnidirectional transducer. Focusing through random media or inside chaotic reverberant cavities is shown to be possible when

broadband pulses are used. A comparison with phase-conjugated experiments shows that this result is not obtained for monochromatic waves. It will be shown that the time-reversal focusing quality depends on the pulse bandwidth and on the statistical decorrelation of the wave fields for different frequencies. A set of experiments conducted through multiple scattering media of different thickness and through closed cavities will be presented. Self-averaging properties of time-reversed fields will be discussed for these various configurations.

FRIDAY MORNING, 8 DECEMBER 2000 CALIFORNIA SALONS 1 AND 2, 9:00 TO 11:15 A.M.

Session 5aPP

Psychological and Physiological Acoustics: Complex Sounds: Physiology to Application

David R. Perrott, Chair
Psychology Department, California State University, University Drive, Los Angeles, California 90032

Contributed Papers

9:00

5aPP1. Context-dependent neural coding in the chinchilla cochlear nucleus. Lori L. Holt (Dept. of Psych., Carnegie Mellon Univ., 5000 Forbes Ave., Pittsburgh, PA 15213), Valerie Ventura (Carnegie Mellon Univ., Pittsburgh, PA 15213), William Rhode (Univ. of Wisconsin—Madison, Madison, WI 53706), Sam Behesta, and Alessandro Rinaldo (Carnegie Mellon Univ., Pittsburgh, PA 15213)

One of the most challenging dilemmas for theories of speech perception is the lack of invariance between acoustic signal and perception. Due to physical constraints upon articulators, there is a good deal of context-dependency both in speech production and in the resulting acoustic speech signal. Consequently, the acoustic pattern most closely related to a given speech sound varies dramatically depending on context. Yet, by some means, the perceptual system perceives these unique acoustic events as linguistically equivalent. This phenomenon is observed experimentally as perceptual context effects whereby adjacent speech can modulate perceived identity of a given speech sound. Recent perceptual results suggest that this phenomenon may be governed by general auditory mechanisms [Holt *et al.*, 1996; Lotto *et al.*, 1997; Lotto and Kluender, 1998; Holt, 1999] rather than speech-specific processes. In the present study, we sought to explore how such context-dependencies might be encoded. We recorded responses of ventral cochlear nucleus (VCN) neurons of anesthetized chinchillas to nonspeech stimulus targets with adjacent context stimuli that varied in spectral content. Results demonstrate context-dependent effects of frequency and intensity.

9:15

5aPP2. Magnetic brain responses to second formant transitions in dependency of transition time, signal voicing, and attention. Ingo Hertrich, Klaus Mathiak, Hermann Ackermann (Neurology Dept., Univ. of Tuebingen, Hoppe-Seyler-Str. 3, D-72076 Tuebingen, Germany, ingo.hertrich@uni-tuebingen.de), and Werner Lutzenberger (MEG-Ctr., Univ. of Tuebingen, D-72076 Tuebingen, Germany)

Using an oddball paradigm, magnetoencephalographic (MEG) recordings were performed during perception of synthetic /i/-like signals (frequent stimuli) and interspersed /bi/ or /yi/-like deviants with an initial second formant ($F2$) transition of 20, 40, 60, or 80 ms. The experiment was run under two different signal conditions: formant-synthesized voiced stimuli or spectrally and loudness-matched unvoiced cognates generated by filtered white noise. Listeners' attention was directed either to the infrequent stimuli or to a moving geometric pattern (visual distraction task). Left- and right-hemispheric M50, M100, and MMNm dipole strengths were derived from the MEG measurements. The results show a strong bilateral effect of signal voicing on the M50 and M100 responses, irrespective of $F2$ transition time, being enhanced with attention directed to the acoustic signal. The deviant stimuli evoked a significant mismatch reaction which under directed attention was lateralized to the left hemisphere. The strength of the MMNm increased across transition time, but exhibited a nonlinearity at 40 ms in terms of an interaction with signal voicing. In conclusion, these results suggest signal-related bilateral preprocessing during the M50 and M100 responses and attention-dependent lateralization of mismatch reactions to speech-relevant acoustic features.

9:30

5aPP3. The relation between backward masking and measures of reading, language, and other cognitive abilities in fourth-grade children. Charles S. Watson, Gary R. Kidd (Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405, watson@indiana.edu), and David A. Eddins (Univ. of Buffalo, Buffalo, NY 14214)

A backward masking test was administered to 247 fourth-grade children participating in the Benton-IU Project, an epidemiological, longitudinal study of the factors that lead to success or failure in grade school. It has been hypothesized that children with language disorders, including reading disabilities, demonstrate high levels of backward masking. Low correlations between masked thresholds, speech recognition abilities, and reading achievement in the overall population provided only weak support for this hypothesis. Within the subgroup of poor readers, there was a substantially larger correlation between masking and speech processing abilities, but neither of these abilities was strongly associated with reading achievement. Reading achievement was much more strongly associated with other cognitive, intellectual, and linguistic measures; the correlation between predicted and obtained teacher-assigned reading grades was 0.73. This study illustrates that research with selected samples of normal and impaired subjects can lead to different conclusions from those based on the entire population. [Work supported by NIH/NIDCD and Indiana University.]

9:45

5aPP4. Individual differences in auditory abilities among normal-hearing listeners. Gary R. Kidd, Charles S. Watson, and Brian Gygi (Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405, kidd@indiana.edu)

An extended version of the Test of Basic Auditory Capabilities (TBAC) [Watson *et al.*, J. Acoust. Soc. Am. Suppl. 1 **71**, S73 (1982)] includes 19 subtests that evaluate listeners' abilities to discriminate and identify auditory stimuli. Stimuli include single tones, tonal sequences, SAM and rippled noise, temporal gaps, nonwords, words, sentences, and environmental sounds. A total of 340 college-student subjects were tested with this battery of tests. A principal components analysis yielded a four-factor solution that accounts for roughly 50% of the variance. The first factor primarily reflects loudness and duration discrimination, the second is associated with sensitivity to temporal envelope variation (SAM noises), the third is associated with identification of highly familiar sounds,

whether they be speech or nonspeech, and the fourth factor includes pitch-discrimination and spectral-temporal tasks, suggesting a common ability to discriminate complex patterns. As in earlier studies in this series, measures of spectral and temporal resolving power for laboratory-generated sounds are very weak predictors of individual differences in speech processing. SAT scores and grades in college courses were also examined to assess the relation between auditory perceptual abilities and abilities in language and mathematics. [Work supported by NIH/NIDCD.]

10:00

5aPP5. The role of timbre similarity in the binding of musical features. Michael D. Hall, Kimberly Wieberg, and James Plonowski (Psych. Dept., Univ. of Nevada Las Vegas, 4505 Maryland Pkwy., Box 455030, Las Vegas, NV 89154-5030, hallm@nevada.edu)

The visual search literature indicates that features are registered separately, and that a feature binding process is required to perceive objects; primary evidence for this process is the occurrence of illusory conjunctions, the perception of incorrect combinations of presented features (e.g., Treisman and Schmidt, 1982). A corresponding auditory process has been established through illusory conjunctions of musical features, though the factors that contribute to auditory illusory conjunctions remain unclear. The role of one potential factor, the perceptual similarity of features, was examined in two experiments with musically trained listeners. In experiment 1, listeners searched pairs of simultaneous, lateralized tones to identify a presented timbre from a target set and its pitch. The perceptual similarity of each target timbre to its nontarget timbre varied across trials. The results were submitted to models of search performance to assess the incidence of illusory conjunctions and feature misperception as a function of timbre similarity and to account for guessing strategies. Experiment 2 directly determined the similarity of the timbres. Perceptual distance was evaluated by submitting the mean ratings of perceived similarity for each pair of timbres to multidimensional scaling analyses. Implications for the nature of auditory feature integration will be discussed.

10:15–10:30 Break

10:30

5aPP6. Effects of low-pass and high-pass filtering on the identification of environmental sounds. Brian Gygi (Psych. Dept., Indiana Univ., Bloomington, IN 47405, bgygi@indiana.edu), Charles Watson, and Gary Kidd (Indiana Univ., Bloomington, IN)

Listeners' abilities to identify environmental sounds under high-pass and low-pass filtering was determined using a constant-stimulus method. The goal was to determine the degree to which listeners rely on information in various frequency regions when attempting to identify naturally occurring sounds. In experiment 1, 70 environmental sounds were low- and high-pass filtered with cutoff values ranging from 300 to 4800 Hz in octave steps. Subjects identified the sounds after an initial familiarization stage in which all sounds were presented without filtering. With low-pass filtering, performance improved from 57% to nearly 100% correct as the cutoff frequency increased. Performance with high-pass filtering was better, with percent correct identification above 70% for all cutoff values. Importance functions derived from these results were used to develop articulation index-like predictions for the identification of a representative set of environmental sounds. [Work supported by NIH.]

10:45

5aPP7. Visualization and auralization of sound quality. Peter Daniel and Harald Mundt (Neutrik Cortex Instruments, Erzb.-Buchberger-Allee 14, D-93051 Regensburg, Germany, daniel@neutrik-cortex.de)

VIPER, a new tool for the VIsual PERception of sound quality, will be presented. Requirement for the visualization of sound quality is a signal analysis modeling the information processing of the ear. The first step of the signal processing implemented in VIPER calculates an auditory spectrogram by a filter bank adapted to the time- and frequency resolution of the human ear. The second step removes redundant information by extracting time- and frequency contours from the auditory spectrogram in analogy to contours of the visual system. In a third step, contours and/or auditory spectrograms can be resynthesized, confirming that only aurally relevant information was extracted. The visualization by contours allows the intuitive grasp of the important components of a signal. Contributions of parts of a signal to the overall quality can be easily auralized by editing and resynthesizing them. Resynthesis of time contours alone allows, e.g., auralization of impulsive components separately from the tonal components. Further processing of the contours determines tonal parts in form of tracks. Audible differences between two versions of a sound can be visually inspected in VIPER through the help of auditory distance spectrograms. Applications are shown for interior noises of cars and for speech.

11:00

5aPP8. Aircraft noise: Intolerability threshold of residents in a single-event-field study. Joachim Kastka (Institut fuer Arbeits und Sozialmedizin, Univ. of Duesseldorf, Universitaetsstr 1, Duesseldorf, D-40225, Germany) and Michael Faust (Univ. of Duesseldorf)

Environmental aircraft noise of an airpath crossing a residential area is defined by single discrete flight events. The aim of this study was to examine if there is a threshold of intolerability of disturbance for overflight levels, measured in dB(A) Lmax of overflight events. Twenty-five residents living under an airpath at Duesseldorf Airport participated in a field study for 8 days. The only instruction given to the residents, living in their homes under natural conditions, their normal way of life, was to press a button on a small portable computer when an overflight event produced intolerable acute noise disturbance. Within the time period of 8 days, the airport produced 2400 overflights on the route leading over the area, mean Leq day (16 h) was 72 db(A). The noise was measured continuously in 1-s interval periods; the mean Lmax was 83 db(A), and 80% of all events were positioned between 74 and 92 dB(A) Lmax. For each flight event the mean intolerability reponse of the residents was computed; 22% of all events were evaluated as intolerable. A nonlinear positive accelerated function was found to best describe the relationship between single-event Lmax and intolerabilty response of residents to the noise events. Referring to this dose-response function, the threshold of intolerability defined by 10% intolerability responses is given by 74 dB(A) Lmax. Stevens power function and the loudness concept are discussed in the context of the results.

Session 5aSC

Speech Communication: Functional Neuroanatomy and Neurophysiology of Spoken Word Recognition and Speech Perception

Edward T. Auer, Chair
House Ear Institute, 2100 West Third Street, Los Angeles, California 90057

Chair's Introduction—9:00

Invited Papers

9:05

5aSC1. Organization of human auditory cortex. Jean K. Moore (Dept. of Neuroanatomy, House Ear Inst., 2100 W. 3rd St., Los Angeles, CA 90057, jmoore@hei.org)

An essential factor in stimulus processing within the auditory cortex is the horizontal transmission of information from one cortical subdivision to another. The importance of transcortical connections is underscored by the developmental pattern of the human auditory cortex. In ontogeny, maturation occurs in horizontal strata, first in the marginal layer, then in the afferent axons to the deeper cortical layers, and last in the commissural and association axons of the superficial cortical layers. Because direct tract tracing studies are not conducted in humans, information as to the direction and extent of transcortical information flow comes from studies of evoked activity and neuronal projections in subhuman primates. These primate studies have developed the core-belt-parabelt model of organization of auditory cortex, with further transmission from the parabelt region to such higher level processing areas as the dorsolateral frontal cortex and the superior temporal sulcus. This framework of information transmission developed in primates fits well with the results of imaging studies in humans. In this hierarchical system, an issue of interest is the level at which processing of speech stimuli conveyed through the auditory channel converges with processing of speech stimuli received through the visual channel.

9:35

5aSC2. Perception of audiovisual speech stimuli. Mikko Sams (Lab. of Computational Eng., Helsinki Univ. of Technology, P.O. Box 9400, FIN-02015 HUT, Finland)

In face-to-face communication, speech is seen as well as heard. The important role of visual speech becomes evident under noisy conditions. The efficacy of visual speech is also nicely demonstrated by the so-called McGurk effect: when the auditory syllable /aba/ is dubbed on to the video-taped visual articulation of /aga/, most subjects hear a syllable /ada/. Recent neuromagnetic recordings have revealed a specific integration signal generated at or close to the primary auditory cortex. The onset of this signal is 150–250 ms. In a similar vein, fMRl studies have indicated that the primary and/or secondary auditory cortical areas are active during silent lipreading. It was recently demonstrated in our laboratory that the integration is similar when nonwords and real words are used as stimuli, suggesting that the integration is not much influenced by lexical representations and probably occurs at the phonetical-perceptual information processing stage. New data from our laboratory preliminarily suggests that the perception of audiovisual stimuli is influenced by attention. This indicates that the integration is not, at least totally, automatic.

10:05

5aSC3. Time course of word identification and semantic integration in spoken language. Cyma Van Petten, Susan Rubin, Marjorie Parks (Dept. of Psych., Univ. of Arizona, Tucson, AZ 85721, vanpettc@u.arizona.edu), Elena Plante (Univ. of Arizona, Tucson, AZ 85721), and Seana Coulson (UCSD, La Jolla, CA 92093-0515)

Human speech consists of a nearly continuous stream of auditory input, so that the semantic message formed by combinations of words must be analyzed as the input continues. The minimum duration signal necessary to identify a set of words was established via the gating technique: Subjects were asked to identify (or guess) words after hearing only the initial 50, 100, or 150 ms, etc. Results showed that most words were identified before their acoustic offset: Average word duration was 600 ms, but identification accuracy was close to 90% after 350 ms of input. The isolation points established in the gating experiment were compared to the time course of semantic integration evident in event-related brain potentials (ERPs). The gated words were used as congruous and incongruous sentence completions (in their full-duration versions). Differential ERP responses to contextually appropriate and inappropriate words were observed by 200 ms after word onset, before the acoustic signal was sufficient to uniquely identify the words. These results indicate that semantic integration can begin to operate with only partial, incomplete information about word identity. If time allows, the talk will also describe similarities and differences between the semantic processing of words and environmental sounds.

10:35–10:45 Break

10:45

5aSC4. Subsystems in human auditory cortex revealed by functional imaging. Sophie Scott (Inst. of Cognit. Neurosci., 17 Queen Square, London WC1N 3AR, UK, sophie.scott@ucl.ac.uk), Stuart Rosen (UCL, London, UK), Catrin Blank, and Richard Wise (Hammersmith Hospital, London, UK)

This study investigated the neural systems involved in speech perception, controlling for the complexity of the speech signal. Using speech stimuli of varying intelligibility but equivalent structural complexity, functional neuroimaging (PET) was used to investigate the brain activation involved in processing intelligible speech in eight right-handed native English speaking volunteers. A stream of processing in the left temporal lobe was demonstrated, associated with the intelligibility of speech. The left superior temporal sulcus responded to the presence of phonetic information, but the anterior region only responded if the stimulus was also intelligible. This novel observation demonstrates a left-lateralized anterior temporal pathway for speech comprehension. Work in visual neuroscience has identified at least two possible functional subsystems in visual processing, a ventral ''what'' pathway passing from visual cortex to the ventral temporal lobes and a dorsal ''where/how'' pathway passing from visual cortex to the parietal lobes. Work in nonhuman primates has revealed analogous pathways in the auditory system. The left lateralized stream of processing shown in this study may form a verbal ''what'' pathway in human auditory perception. The right temporal lobe, in contrast, showed selective activation for stimuli that contained dynamic pitch variation, regardless of intelligibility.

11:00

5aSC5. The interface between the perception and production of speech. Richard Wise (MRC Cyclotron Unit, Hammersmith Hospital, Du Cane Rd., London W12 0NN, UK, richard.wise@csc.mrc.ac.uk) and Sophie Scott (Inst. of Cognit. Neurosci., London WC1N 3AR, UK)

Four positron emission tomography studies were reanalyzed to identify anatomically separable, functional subsystems associated with speech perception, speech production, hearing complex, nonspeech sounds, and verbal semantic category verbal fluency. The results from a total of 26, right-handed, male volunteers were included. The results indicated two separate caudal (posterior) ''streams'' of auditory processing. One, directed along the supratemporal cortical plane, responded to both nonspeech and speech sounds, including the sound of the speaker's own voice. Activity in response to perception merged at the temporo-parietal junction with a region where activity was linked to speech production. In contrast, the left superior temporal sulcus below (ventrolateral to) the caudal supratemporal plane responded only to an external source of verbal stimuli: but, in addition, it was activated by the recall of lists of words during verbal fluency. The results are compatible with a hypothesis that the caudal superior temporal cortex is specialized for the mimicry of sounds, including repetition, the role of the caudal left superior temporal sulcus being to transiently represent verbal sequences, whether heard or internally generated and rehearsed. These processes are central to the perception, rehearsal (both covert and overt), and eventual acquisition of long-term lexical memories of novel words.

11:15

5aSC6. Neural correlates of duplex perception: A whole-head magnetencephalography study. Klaus Mathiak, Ingo Hertrich, Werner Lutzenberger, and Herrmann Ackermann (Univ. of Tuebingen, MEG Ctr., Otfried-Mueller-Str. 47, D-72076 Tuebingen, Germany, Klaus.Mathiak@Med.Uni-Tuebingen.de)

Simultaneous experience of the same stimulus events in two distinct phenomenological modes is referred to as duplex perception (DP). The most widely investigated DP paradigm decomposes the formant pattern of the stop consonant-vowel (CV) syllables /ga/ and /da/ into two complementary portions: the isolated transient of the third formant (''chirp'') and the remaining sound structure (''base''). Presentation of the two syllable components to opposite ears results in the perception, first, of the chirp and, second, the fused CV syllable. In order to further elucidate the mechanisms underlying DP, the present study recorded mismatch fields in response to a series of dichotically applied base and chirp components using whole-head magnetencephalography (MEG). This magnetic analog to mismatch negativity is considered a correlate of early cognitive stimulus representation. (a) Under a visual distraction task, the preattentive mismatch fields were stronger in response to contralateral deviants and lateralized to the left hemisphere. (b) Under attention toward the fused percept /da/, left ear deviant chirps yielded an *ipsilateral* enhanced and posteriorly shifted reaction. These data suggest representations of the phonetic percepts in secondary dominant auditory cortex. Perception tasks may activate these templates for syllable fusion.

11:30–11:50

Panel Discussion

Session 5aUW

Underwater Acoustics: Matched Field Processing and Beamforming

Gerald L. D'Spain, Chair
Marine Physical Laboratory, Scripps Institution of Oceanography, 291 Rosecrans Street, San Diego, California 92106

Chair's Introduction—8:25

Contributed Papers

8:30

5aUW1. Comparison of wide-band matched field processing algorithms. Paul Hursky, Chris O. Tiemann, Michael B. Porter (Science Applications Intl. Corp., 888 Prospect Ave., Ste. 201, La Jolla, CA 92037), and Richard B. Evans (Sci. Applications Intl. Corp., New London, CT 06320)

There are several ways to realize a wide-band MFP process. The data may be processed in the frequency domain, one frequency at a time, and then incoherently or coherently integrated over the band. Alternatively, the data may be processed in the time domain, by matching modeled correlations (e.g., calculated using a ray trace model) with measured correlations. Auto correlations at a single sensor or interelement cross correlations can be used. We compare the frequency-domain and time-domain approaches with respect to array configuration requirements, sensitivity to array element location errors, and sensitivity to environmental mismatch. We demonstrate our findings using simulated and experiment data.

8:45

5aUW2. Array design and motion effects for matched field processing. Brian Tracey, Nigel Lee, Lisa Zurk, and James Ward (MIT Lincoln Lab., Group 103, 244 Wood St., Lexington, MA 02420-9108)

While matched field processing (MFP) algorithms have most commonly been applied to data collected on vertical line arrays, there is increasing interest in applying MFP to more general array geometries. One challenge for novel array design is to determine the fundamental performance of volumetric arrays in a shallow water channel. To address this issue, recent work by Tantum and Nolte [J. Acoust. Soc. Am. **107**, 2101–2111 (2000)] has been extended to allow more general array geometries, including horizontal arrays steered away from endfire. The correlation among acoustic modes across the array can be related to MFP resolution and sidelobe levels. Another trend in underwater array design is the consideration of larger arrays which provide additional array gain. However, large arrays are susceptible to the effects of motion [Zurk *et al.*, Oceans 99]. A new time series simulation that rigorously treats source and receiver motion is used to quantify the effects on adaptive MFP. Simulation results are compared to data from the Santa Barbara Channel experiment (SBCX). [This work was sponsored by DARPA under Air Force Contract F19628-95-C0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Air Force.]

9:00

5aUW3. Preliminary results of an acoustic barrier test using matched-field tracking. J. Mark Stevenson (Acoust. Branch, Code D857, Space and Naval Warfare Systems Ctr., San Diego, CA 92152) and Homer P. Bucker (Space and Naval Warfare Systems Ctr., San Diego, CA 92152)

In September 2000, a prototype autonomous seafloor barrier HLA (horizontal line array) will be deployed in the Adriatic Sea. This test apparatus is not designed as a beamforming array, but rather contains six hydrophones nonuniformily distributed over 650 m in a line on the seafloor. The test objective is to make automatic detections of acoustic sources traveling over the barrier using matched-field tracking [H. Bucker, J. Acoust. Soc. Am. **96**, 3809–3811 (1994)]. The tracking algorithm operates in real time in a digital signal processor (DSP) running autonomously inside the apparatus. The procedure is to select N Fourier frequency bins based on S/N ratios after CFAR normalization, and then to determine the best tracks using M-of-N processing. Results will be shown comparing real-time results with those obtained in postexperiment analysis. [Work supported by ONR, Code 321-SS.]

9:15

5aUW4. Robust adaptive nulling in matched field processing. J. S. Kim, W. A. Kuperman, H. C. Song, and W. S. Hodgkiss (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0238, jkim@mpl.ucsd.edu)

Nulling out sources of interference is a major goal of adaptive array processing. For moving interferers, null-broadening is a useful technique to reduce the interference. A null-broadening approach has been recently developed for planewave beamforming [R. J. Mailloux, Electron. Lett. **31**, 771–772 (1995)] [M. Zatman, *ibid.* **31**, 2141–2142 (1995)] to cover the angular extent of the motion. One way to achieve this is to introduce the effects of dispersion into a covariance matrix, which is equivalent to averaging the covariance matrix over a finite bandwidth in frequency. In this study, the null-broadening approach is extended to the matched field processing (MFP) in an ocean waveguide, based on an MV processor with a white noise constraint. In particular, the null-broadening in range is accomplished through using waveguide invariant properties. Simulations as well as the SWELLEX-96 experimental data clearly demonstrate the null-broadening effect.

9:30

5aUW5. Incoherent beamforming with interference patterns. Gerald L. D'Spain, Lewis P. Berger, William S. Hodgkiss, William A. Kuperman, and Fred N. Spiess (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0704)

A conventional coherent plane wave beamformer exploits the phase evolution of the wave field across the array aperture, using the phase at one element as reference, to obtain information on source location. Alternatively, in a multipath environment, the phase of one mode or multipath can be used as reference for the phase evolution of the other multipaths in the wave field. This ''intermode interferometry'' is the basis of an incoherent beamformer presented here. It involves summing together individual element autospectra, after performing a frequency-domain dilatation/contraction operation based upon a hypothesized source location. Because it uses autospectra rather than element time series, this processing provides no gain of signal over background noise. Therefore, its usefulness in its basic form is restricted to localization of strong signals. However, it is not limited to line arrays with equal element spacing, and the array element positions need only be known to a fraction of an interference wavelength rather than an acoustic wavelength. Examples of the

incoherent beamforming results from simulation and actual ocean acoustic data collected during the shallow-water (200 m) SWellEx-3 experiment will be presented. Emphasis is placed on 2D passive synthetic aperture applications. [Work supported by ONR, Code 321US.]

9:45

5aUW6. Source localization using the waveguide invariant approach and conventional plane wave beamforming with vertical hydrophone array data. Gerald L. D'Spain, William A. Kuperman, and James J. Murray (Marine Physical Lab., Scripps Inst. of Oceanogr., La Jolla, CA 92093-0704)

Broadband source localization, both in range and depth, can be performed using conventional plane wave beamforming with vertical hydrophone array data. Little computation or prior knowledge of the environment is required. An analytical model, based on the application of the waveguide invariant approach to a normal mode formulation of the sound field, is developed to predict the behavior of striation patterns as a function of vertical wave number, range, and frequency. The approach is illustrated through numerical simulation and applied to data collected during the shallow-water (200-m) SWellEx-3 experiment in 1994. Extension to canonical deep-water environments also is discussed. [Work supported by ONR, Code 321US.]

10:00–10:15 Break

10:15

5aUW7. Acoustic tomography for optimizing matched field processing on a tracking range. Colin J. Lazauski, Joan M. Cembrola (Naval Undersea Warfare Ctr. Div. Newport, 1176 Howell St., Newport, RI 02841, lazauskicj@npt.nuwc.navy.mil), and James H. Miller (Univ. of Rhode Island, Narragansett, RI 02882)

A future Navy passive tracking range will use a real-time matched field processing (MFP) method. As environmental uncertainty affects the performance of matched field processing (MFP) schemes, a need for near real-time 3-D sound speed profile (SSP) estimations exists. The Navy MFP tracking range will be equipped with a tomographic system to obtain these estimates. This will reduce the error risk of MFP by having timely *in situ* data available. The tracking range considered covers on the order of up to 340 km^2 and is located in shallow water with a sloping bottom (200–600 m). Time and space scales of the oceanography can be short in these regions and rapidly degrade the MFP solutions. Obtaining a real-time estimation of the sound speed field using conventional means requires more time and facilities than are available and therefore the use of tomographic methods is proposed. This study shows that the acoustic tomography system designed for the Navy range can achieve an accuracy of better than 1 m/s. Spatial resolution of 4 km in the horizontal and 10 m in the vertical are possible with eight sources and four vertical receiving arrays.

10:30

5aUW8. Source localization in a shallow water random environment. Catherine Stamoulis (Dept. of Ocean Eng., MIT, 77 Massachusetts Ave., Cambridge, MA 02139 and Naval Res. Lab., 4555 Overlook Ave SW, Washington, DC 20375), Marshall Orr (Naval Res. Lab., Washington, DC 20375), and Ira Dyer (MIT, Cambridge, MA 02139)

Source localization in a random ocean environment has been investigated through analysis of data from the SWARM 95 experiment. Both broadband matched-field processing and matched-mode processing have been used for localization. In particular, acoustic data collected with a vertical line array and a moving upsweep source have been analyzed, in the frequency range 275–350 Hz. The source distance varied between 2.5 and 27 km. Due to the dynamic nature of the shallow water column and in particular the presence of soliton packets, the available environmental data used in matched-field processing was not sufficiently accurate for correctly localizing the source via this method. The mismatch between data and replica fields, particularly their phase differences, was significant. Source range errors were of the order of the source distance from the array, 2 to 20 km. In contrast, matched-mode processing yielded much better results. The modal time series were first estimated from the data, mode filtering was then used to eliminate noisy or nonparticipating modes and localization was finally performed. The source range errors varied between 200 m to 5 km.

10:45

5aUW9. Rank-*N* stochastic matched field processing. Peter M. Daly and Arthur B. Baggeroer (MIT, Cambridge, MA 02139, pmd@mit.edu)

Underwater acoustic source localization can be performed using matched field processing (MFP). Successful localization depends on accurate signal and environmental input parameters to a propagation model. Some parameters, notably the sound velocity profile (SVP), vary both spatially and temporally, and are thus difficult to treat as deterministic quantities. Instead, one can model the SVP and the resultant pressure field as random vectors. This allows one to detect targets using second order statistics of the received and modeled pressure fields. Baggeroer *et al.* [J. Acoust. Soc. Am. **106**, 2126] demonstrated successful target localization using individual singular vectors of a modeled covariance matrix. A new method of MFP is proposed which combines energy from orthogonal singular vectors of the modeled covariance matrix. Each correlation is weighted and summed to yield a detection statistic. A maximum-likelihood nonrandom parameter estimation method is then applied to determine the source location. Through inclusion of signal energy which would normally be lost to environmental mismatch, this method significantly reduces estimation bias in a simulated 200-m shallow water waveguide. Localization results are shown using both simulated and experimental data from the 1998 Santa Barbara Channel Experiment. [Work supported by SAIC, DARPA and ONR.]

11:00

5aUW10. A Gibbs sampling approach to matched-field source localization and deconvolution. Zoi-Heleni Michalopoulou (Dept. of Mathematical Sci., New Jersey Inst. of Technol., Newark, NJ 07102)

In narrow-band source localization the source spectrum is typically unknown. Matched-field techniques employ normalization schemes to remove the uncertainty for the unknown amplitude at the source, whereas consideration of phase differences between replicas and true field addresses the unknown phase issue. In this work we implement a matched-field technique that simultaneously estimates source location and source spectrum. The estimation is performed using an efficient Gibbs sampling scheme. Results of the Gibbs sampling estimation approach are compared to those of the linear (Bartlett) processor under several noise levels. The comparison reveals that the simultaneous process of deconvolution and localization has superior performance to that of the linear localization process. Moreover, improvement in localization is obtained without a significant increase in computational requirements. In addition to the improved performance, the proposed method has the advantage of providing more information than conventional matched-field localization, since it also estimates the source spectrum which is often of great interest. [Work supported by ONR.]

11:15

5aUW11. Stabilizing Rietschs signal estimation technique against noise for multipath mitigation in shallow water transient classification. George B. Smith (Naval Res. Lab., Code 7185, Stennis Space Center, MS 39529-5004) and James P. Larue (Univ. of New Orleans, Lakefront, New Orleans, LA)

Extraction of a transient signal received on a hydrophone array at tactical ranges in shallow water is made difficult by distortions caused by dynamic multipath. Blind deconvolution techniques can adapt the processing to the environment. Rietschs algorithm [E. Rietsch, Geophysics **62**(6), 1931–1946 (1997)], which is an exact algebraic solution to the blind deconvolution problem, performs exceptionally well when no noise is

present in the data. In this work, progress with attempts to stabilize Rietschs algorithm using noisy simulated data is presented. [This work was supported by the Office of Naval Research and the Naval Research Laboratory.]

FRIDAY AFTERNOON, 8 DECEMBER 2000 CATAMARAN/TRIMARAN ROOMS, 1:30 TO 4:10 P.M.

Session 5pAA

Architectural Acoustics: Measures of Auditorium Acoustics

Angelo J. Campanella, Chair
Campanella Associates, 3201 Ridgewood Drive, Hilliard, Ohio 43026-2453

Contributed Papers

1:30

5pAA1. Subjective threshold of sound strength (G) due to level changes in the early sound field: Preliminary results. Jerald R. Hyde (JRH Acoustics, Box 55, St. Helena, CA 94574)

Acoustical measurements and related subjective studies over the past two decades have shown that increased early reflections produce increased sound strength, G in concert auditoria. The temporal distribution of early reflected sound and the role of room geometry have been found to be important design related factors relating to both G and acoustical intimacy, and therefore to hall quality in general [J. R. Hyde, Proc. 16th ICA, Seattle, Paper 1aAA3 (1998)]. A listening test was devised to determine the threshold of perception of changes in loudness relative to changes in the level of the early sound field. The test setup consisted of direct and single-delay speaker channels in a semi-anechoic room, using brief samples of different styles of symphonic music. Subjects were asked to choose the louder of two paired samples of the same musical passage, played in rapid succession. Preliminary results from ''unsophisticated'' test subjects indicate that the perception threshold of loudness difference occurs when the total sound strength G varies by 1.0 dB due to the addition of a single early reflection at about -6 dB. Details and results of the experiment will be presented, and ideas for future work and test procedures will be solicited.

1:45

5pAA2. Subjective preference of initial time delay gap in a Javanese gamelan concert hall. J. Sarwono and Y. W. Lam (School of Acoust. and Electron. Eng., Univ. of Salford, Brindley site, Salford M7 9NU, UK)

Several human subjects were involved in a subjective test using the paired comparison method to determine the preferred initial time delay gap (ITDG) for listening to Javanese gamelan music (gendhing) in a concert hall. To provide comparisons, acoustic measurements in several pendopos in Indonesia have also been carried out. Pendopo is an open-sided hall, where the Javanese gamelan is usually played. The measurements were conducted using the room impulse response method, with a maximum length sequences (MLS) signal as its sound source. The ITDG of four pendopos are presented in this paper. The results of subjective preference testing showed that there are preferences for short ITDG, although evidence for a single value of preference was not conclusive. These results agree with the ITDG from the room responses measured in the pendopos, which are not common concert halls but open-sided halls.

2:00

5pAA3. Subjective assessment of stage environments for chamber music. Weichun Wang (Saint Mary and Saint Jones Inst. of Technol., 499, section 4, Danchin Rd., Dansui, Taipei County, Taiwan, ROC, VGN@tpts6.seed.net.tw), Weihwa Chiang, and Shih Tang Chen (Taiwan Univ. of Sci. and Technol., 43, Taipei City, Taiwan 106, ROC)

Stage acoustics were evaluated in four halls regarding chamber music performances. Sensations about hearing oneself, hearing others, ease of ensemble, and overall impression were expressed by players of violin and piano sonata, piano trio, and brass quintet. Evaluations were taken at two locations with and without detached reflectors in each hall. Subjects in the audience were asked about the balance among players. In additional to the well-accepted support factor (ST1), calculated measures included a similar parameter early-to-direct energy ratio (ED80), where the energy integrated from 20 to 100 ms was substituted by the energy from 7 to 80 ms since recital halls were normally smaller than symphony halls. As a result, most acoustical environments were acceptable to the players. Overall impression was significantly correlated with all of the subjective attributes and there was a strong relation between hearing others and ease of ensemble ($r=0.921$). The players of brass quintet favored higher ED80 values than did the others. For the audience, significant variances of perceived strength were found in cello for piano trio and in horn for brass quintet. On-going research is conducted to collect data in other halls. [Work supported by the National Science Council of Taiwan.]

2:15

5pAA4. Measurements of stage support on a 1/25 scale model of an orchestra platform. Jin Yong Jeon (School of Architect. Eng., Hanyang Univ., Seoul 133-791, Korea, jyjeon@email.hanyang.ac.kr)

Stage support values known as ST1 and ST2 were measured in different stage positions both in a 1/25 scale model and a computer model of an actual large concert hall stage. Tests on the scale model were made both with and without model musicians. For the measurement of ST, different distances to the first musician were tried; 2m clear space around the source and receiver looks optimal. The effect of the side wall profile of the stage was investigated; the profile in the actual hall was found to be unhelpful for support. Different positions were also tested for the orchestra on the large stage. The position with greatest acoustic support was found to be at the rear of the stage, not the position close to the front of the stage, which is commonly used.

2:30

5pAA5. The effects of cylindrical diffusers on sound diffusion in a rectangular concert hall. Weihwa Chiang, Hueiping Wu, and Yankun Shu (Taiwan Univ. of Sci. and Tech. 43, section 4, Kee-lung Rd., Taipei City, Taiwan 106, ROC, WHCH@mail.ntust.edu.tw)

Acoustical measurements in a 1:20 scale model were conducted to investigate the influence of a 1/4-cylinder diffuser on sound diffusion in a rectangular concert hall. Diffusion was evaluated regarding temporal and directional distribution of acoustical energy at individual seats as well as energy distribution over all seats. With volume of 19,000 m3 and seating area of 1,200 m2, the size and shape of the hall were determined based on three classical rectangular concert halls. Overall, diffusers around the stage were more effective for the source at the center, front of the stage than the source at the side, rear of the stage. The hall average values of late interaural cross correlation (IACCL3) were reduced by using transversal diffusers on ceiling or longitudinal diffusers on the side walls. Although the influences on the hall average values of echo criteria (EC) and temporal diffusion (TD) were relatively small, diffusers on the rear stage wall were effective in reducing EC values for the seats near the side walls. Strength factor (G) at the seats under the balcony was significantly increased by placing diffusers on the ceiling. On-going research is conducted to evaluate the effects of diffusers in the audience section. [Work supported by the National Science Council of Taiwan.]

2:45–2:55 Break

2:55

5pAA6. Modal response of coupled spaces. Dorothy A. Najolia, Anthony A. Atchley, Courtney B. Burroughs, and John Fahnline (Grad. Prog. in Acoust., Penn State Univ., State College, PA 16804)

To simulate coupled rooms, two rectangular enclosures were built with a common removable partition between them. The frequency response and spatial distribution of the sound fields were measured when the spaces were coupled together. The inner dimensions of the volumes are 43-by-33-by-26 in. and 32-by-25-by-20 in. The walls are constructed from 0.75-in-thick medium density fiber board. Each enclosed space was instrumented with a 2-D microphone array that was moved in the third dimension, so that the sound field was measured over a 3-D grid. The array in the larger enclosure consisted of 70 electret microphones arranged on a 4-by-4-in. grid, while the array in the smaller enclosure consisted of 42 microphones arranged on a 3.5-by-4-in. grid. An externally coupled compression driver located in a corner excited the larger enclosure. The common wall was interchangeable so that location and size of the opening between the enclosures could be studied. A detailed description of the enclosures is presented, as well as representative data showing the utility of the enclosures in investigating the effects different couplings have on the sound fields. [This work is supported in part by the Penn State Applied Research Laboratory and the College of Engineering.]

3:10

5pAA7. The color of early sound arrivals in an auditorium. U. Peter Svensson (Dept. of Telecommunications, Norwegian Univ. of Sci. and Technol., NO-7491 Trondheim, Norway), Rendell R. Torres (Chalmers Univ. of Technol., SE-41296 Göteborg, Sweden), and Herman Medwin (Ocean Acoustics Assoc., Pebble Beach, CA 93953)

In computational room acoustics, most present-day methods calculate specular reflections complemented by various models for surface scattering. A recent method of edge diffraction calculation [U. P. Svensson *et al.*, J. Acoust. Soc. Am. **106**, 2331–2344 (1999)], builds on the BTM technique which has been proven to be accurate in such diverse situations as noise control by barriers and sound scatter from the rough ocean surface. Here, it is applied to a large concert hall. The hall chosen is one of those in an International Round Robin study. It is shown that calculations of edge diffraction give quantitative details of what has until now been approximated and loosely called ''scattering.'' Here, the early part of the impulse response is considered in order to study aspects such as the frequency content of individual early diffractions and reflections and the low-frequency energy distibution in the room. Only rigid surfaces, and up to second-order edge diffraction and/or specular reflections are taken into account, but the possible extension towards realistic wall impedances and higher-order reflections are discussed. Since the BTM technique is exact and deterministic rather than approximate and statistical, it provides a secure basis for predictions in room acoustical modeling.

3:25

5pAA8. Comparison of acoustical measurements, calculations, and software for a seminar room. Jerry P. Christoff (Veneklasen Assoc., 1711 16th St., Santa Monica, CA 90404, jchristoff@veneklasen-assoc.com) and Rebeka Vital (Dept. of Architecture, USC)

The Verle Annis Gallery is a rectangular room with a volume of approximately 20 000 ft^3. Used primarily for seminars, it is known for its poor acoustic qualities. Various acoustical parameters were measured in the space using MLSSA software. The results were compared to calculations based upon the room geometry and the acoustical properties of the room surfaces. Finally, CATT-Acoustic software was used to model the room in its present condition and also with the addition of acoustical treatment and sound reflecting panels. A novel mechanized device was designed to improve the sound isolation between two simultaneous presentations. A comparison of the measured, calculated, and simulated results are presented. [This work was partially supported by the Paul S. Veneklasen Research Foundation.]

3:40

5pAA9. Case studies of two performing arts halls. David A. Conant and Thomas J. McGraw (McKay Conant Brook, Inc., 5655 Lindero Canyon Rd., Ste. 325, Westlake Village, CA 91362)

A new 1600-seat multipurpose hall in Mesa, Arizona and a 1200-seat music-only concert hall in Pocatello, Idaho were recently designed employing CATT-Acoustic and laser pointer modeling techniques to optimize room shapes and surface properties. Unique challenges included the correct modeling of a multitude of singly curved and doubly curved room surfaces in critical upper volumes. Acoustical design results of the halls, especially with regard to the CATT-Acoustic computer models, will be presented.

3:55

5pAA10. Acoustical characteristics of old wooden churches in Serbia. Miomir Mijic and Dragana Sumarac-Pavlovic (Lab. of Acoust., Faculty of Elec. Eng., Bulevar kralja Aleksandra 73, 11000 Belgrade, Yugoslavia)

There are more than 50 churches in the central part of Serbia built entirely of wood. These churches were all built at the beginning of the 19th century. This paper presents the results of acoustical measurements conducted in these places of worship. In addition, the analysis of acoustical properties of different architectural styles and constructions within the same category of churches is presented.

Session 5pPA

Physical Acoustics: Outdoor Sound, Mostly

Murray S. Korman, Chair
Physics Department, U.S. Naval Academy, Annapolis, Maryland 21402

Contributed Papers

1:15

5pPA1. Buried objects detection and classification with a 3-D parametric sonar. Stéphane Guyonic (DCE/GESMA, BP 42, 29240 Brest-Naval, France, guyonic@gesma.fr)

This paper deals with a new sonar concept suitable for mine detection and classification in shallow and very shallow water. This concept is based on a 3-D high-resolution imaging information obtained from a 2-D scanning. Sediment penetration and narrow-beam constraint is made possible by the use of the parametric technique. A specific tool has been designed to be able to make a connection between a real operational configuration and experimental needs. A 2×2 m^2 suspension frame has been built to allow the sonar displacement along the two axis direction. This suspension frame has been deployed first in a pool and then at sea above three different areas where mines and rocks had been buried. Experiments have been conducted in a shallow water area with the suspension frame mounted 2 m above the sea floor made of rough sand. Three-dimensional imaging techniques have been used to process the data and very good results have been obtained. Underwater mine and object detection and shape classification have been successfully demonstrated up to half a meter sediment burial. This paper shows these first results and describes sonar and signal characteristics and then discusses the data and image processing.

1:30

5pPA2. Nonlinear seismoacoustic land mine detection and discrimination technique. Dimitri Donskoy, Alexander Ekimov, and Jianhui He (Stevens Inst. of Technol., Hoboken, NJ 07030)

The paper presents further development [D. Donskoy, Proc. SPIE **3392** (1998); *ibid.* **3710** (1999)] of the innovative seismoacoustic technique for detection and discrimination of artificial objects, such as mines, pipes, containers, etc., buried in the ground. The present paper discusses a physical model to explain land mine detection and discrimination using acoustic waves and presents results of laboratory and field tests validating the model. At low frequencies (below 1 kHz) the model treats the soil–mine system as a lumped-element system, which includes the mass and stiffness of the soil on top of the mine and the stiffness of a mine diaphragm as the governing parameters. The model explains both: linear and nonlinear dynamic behavior of buried mines, observed in numerous laboratories and field tests. Thus, the developed model helps to evaluate the mine detection capabilities of seismoacoustic techniques based on mine type, burial depth, and soil condition.

1:45

5pPA3. Measurements of the induced pore pressure in packings of glass beads due to an incident airborne acoustic wave. Craig J. Hickey, James M. Sabatier (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677, chickey@olemiss.edu), and William E. Goddard (Harding Univ., Searcy, AR 72149)

The Biot theory predicts the existence of two *P* waves (referred to here as type I and type II) and one *S* wave. In a previous experiment where airborne sound was incident upon an air-filled packing of sand, it was postulated that the signal detected by microphones at shallow depths was due to the type II *P* wave whereas at the deeper depths the signal is associated with the type I *P* wave [C. J. Hickey and J. M. Sabatier, J. Acoust. Soc. Am. **102**, 128–136 (1997)]. In this presentation we discuss the results of measurements obtained in packings of glass beads. Four different packing were respectively made with 304 (30–40 mm), 405 (40–50 mm), 507 (50–70 mm) and 1014 (100–140 mm) size glass beads. A smaller bead size increases the flow resistivity of the packing, thereby increasing the attenuation of the type II *P* wave. If the previous interpretation is correct then the depth at which the signal is dominated by type I or type II *P* wave should change with glass bead size.

2:00

5pPA4. A benchmark fast-field program model for infrasound propagation. Carrick L. Talmadge and Kenneth E. Gilbert (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677)

Three decades ago Pierce [e.g., Pierce, Posey, and Iliff, J. Geophys. Res. **76**, 5025 (1971)] developed an exact theory for acoustic wave propagation in a temperature- and wind-stratified atmosphere. The theory includes both the acousto-gravity wave component (i.e., the Lamb wave) and the effect of a vector wind that can vary with height in both magnitude and direction. In this paper, a fast-field program (FFP) is discussed which numerically integrates Pierce's differential equations, thus eliminating the need to approximate the solutions with a normal-mode expansion. For a stratified atmosphere, the FFP solution can serve as a benchmark for testing parabolic equation (PE) solutions that include the Lamb wave and a vector wind. Pierce's theory is outlined and a general numerical approach is given for solving the fundamental equations of the theory. The numerical solutions are compared with known solutions obtained by other methods, such as the parabolic equation and analytic models. [Work supported by the U.S. Army Space and Missile Defense Command.]

2:15

5pPA5. The Lamb wave in parabolic equation models. Kenneth E. Gilbert (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677) and Xiao Di (Appl. Res. Lab., Penn State Univ., State College, PA 16804)

The inclusion of the Lamb wave (i.e., the inclusion of gravity) in two existing atmospheric parabolic equation (PE) models is discussed. The two models are based on, respectively, the Green's function parabolic equation (GFPE) of Gilbert and Di and the split-step pade parabolic equation (SSPPE) of Collins. While the SSPPE results essentially duplicate recent research by Lingevitch *et al.* [J. Acoust. Soc. Am. **105**, 3049–3056 (1999)], the GFPE research is novel and gives physical insight into the nature of the Lamb wave. Each of the two PE models considered has its own particular advantages and limitations. The GFPE, for example, has an explicit term for the Lamb wave so that physical insight is enhanced. However, including gravity in the GFPE requires an assumption of an exponentially stratified density, an approximation that is valid only to a height of 50–100 km. The SSPPE has the advantage of treating gravity exactly for an arbitrary density stratification. Unfortunately, separating the Lamb wave contribution from the total acoustic field is not straightforward with the SSPPE. This paper outlines the theories for the GFPE and SSPPE and presents numerical comparisons between the two models for a number of infrasound propagation scenarios. [Work supported by the U.S. Army Space and Missile Defense Command.]

2:30

5pPA6. Acoustic pulse propagation over laterally inhomogeneous ground. Donald G. Albert (ERDC-CRREL, 72 Lyme Rd., Hanover, NH 03755, dalbert@crrel.usace.army.mil)

Measurements were conducted outdoors over paths having areas with vastly different ground properties to investigate the effect of acoustic wave interaction with inhomogeneous ground conditions. Acoustic pulses with a bandwidth of approximately 10 to 500 Hz were produced by blank pistol shots, and were digitally recorded after propagating horizontal distances of 30 to 150 m. A natural snow cover was used to control the ground properties. The first measurement was conducted over an undisturbed snow cover, a highly porous material with low acoustic impedance. Then, portions of the snow cover were removed by plowing and the measurements repeated. By removing the snow, a less porous, higher impedance frozen ground surface was introduced into the propagation path. The snow cover was removed in stages so that several different inhomogeneous ground conditions were sampled, and a final measurement was made with the snow cover entirely removed. Changes in the pulse waveforms were observed when only a small proportion of the propagation path was modified by plowing. The measurement results will be compared with existing theories. [Work supported by U.S. Army.]

2:45

5pPA7. Low frequency sound propagation across a land-water boundary. Michael J. White (Applied Research Assoc., 5941 S. Middlefield Rd., Littleton, CO 80123)

Sound waves propagating beyond a shoreline experience changes in surface impedance, surface roughness and refraction profile with distance. Measurements of the temperature and wind with tethersonde balloons indicate significantly weaker sound speed gradients over water than over land. At times when over land a temperature inversion was observed, the inversion was always smaller over the water. When the solar heating produced strong upward refraction near the land surface, the temperature profile over water was neutral (lapse). Simulations were run with a parabolic equation model incorporating range dependent boundary refraction profiles and range dependent boundary conditions. The signal levels exhibit changes with distance due to strong refraction over land, but the changes nearly cease in places where the propagation extends over water. Measured spectrum levels from explosion sound propagating across land-water boundaries agree well with the simulations. [Work supported by USACERL, CHPPM and NAVFAC.]

3:00

5pPA8. Wind gusts or scattering by turbulence: Which is more important for near-ground propagation? D. Keith Wilson (U.S. Army Res. Lab., AMSRL-CI-EP, 2800 Powder Mill Rd., Adelphi, MD 20783-1197, dkwilson@arl.army.mil)

''Wind gusts'' are actually just very large scale turbulence, generally having spatial dimensions comparable to the thickness of the atmospheric boundary layer (several hundred to several thousand meters). For propagation paths smaller than the dimensions of the wind gusts it is conceptually useful to distinguish between variations in the vertical profiles of wind and temperature along the propagation path (caused by the gusts) and scattering by smaller, wavelength-scale turbulent eddies. For example, in the case of upwind propagation, the gusts can cause the boundary of a refractive shadow zone to shift in position, whereas the smaller-scale eddies are responsible for scattering sound energy into the shadow. In this paper a numerical propagation model is used to closely examine the relative roles of the gusts and smaller-scale eddies for upwind, crosswind, and downwind propagation. The strength of both the gusts and smaller-scale eddies is modeled with turbulence similarity theory. The gusts are generally most important for low frequencies (below 100 to 200 Hz), but scattering by wavelength-scale eddies dominates at higher frequencies.

3:15

5pPA9. Acoustical scattering amplitudes of waveletlike atmospheric temperature distributions. Harry J. Auvermann (Computational & Information Sci. Directorate, U.S. Army Res. Lab., 2800 Powder Mill Rd., Adelphi, MD 20783-1197, hauverma@arl.mil)

As an alternative to Fourier space representation of turbulence, wavelet decomposition has certain advantages associated with the ability to calculate effects localized in space, such as turbulent intermittency. Wavelet-based turbulence simulation is particularly useful when calculating the acoustic field scattered into a detector by a turbulent region of the atmosphere. Wavelet decomposition theory utilizes the orthonormal properties of these functions to make a chosen distribution easier to simulate. In this paper, the first step in the process, that is determination of the scattering amplitude, is reported for a number of candidate waveletlike functions. Since a wavelet is defined to have a zero volume integral, some candidates are not wavelets, but may have useful properties. The candidates are: Gaussian, exponential, ''Mexican hat,'' Morlet, cosine, and concentric spheres. While the method of generating three-dimensional velocity distributions is indicated briefly, the results are confined to temperature turbulence. The scattering amplitudes are calculated using the Born approximation in the far field. The concentric sphere candidate allows an exact solution in the form of an infinite series. The result of this calculation is presented also.

3:30

5pPA10. Rocket launch acoustics modeling. Michael J. White and Gary M. Ogg (Applied Research Assoc., 5941 S. Middlefield Rd., Littleton, CO 80123)

The exhaust of a launch vehicle is composed of a high speed jet of gases that interact with the ambient air, the launch pad and the launch structure. This interacting flow is highly turbulent and is a significant source of sound power. Noise from the rocket plume can represent a threat to sensitive components on the vehicle and to the rocket payload (e.g., satellite). Detailed measurements have been made of sound power from several scale model rockets during the lift-off phase. These measurements have been used to successfully predict the sound pressure at other points on the launch pad and on the vehicle with fairly simple relations for propagation, reflection, and diffraction of sound. Work is now underway to automate the calculations, so that designers of future rockets or launch vehicles can assess the noise load in a lift-off environment. Future plans include extension of the model for use in prediction of community noise. [Work supported by NASA and AFRL.]

3:45

5pPA11. Nonlinear scattering of crossed ultrasonic pulsed beams in the presence of turbulence: Doppler shift and spectral broadening effects. Murray S. Korman (Dept. of Phys., U.S. Naval Acad., Annapolis, MD 21402, korman@nadn.navy.mil)

The nonlinear scattering of two finite-amplitude mutually perpendicular crossed beams—interacting in the presence of turbulence—generates a sum frequency component that radiates outside the interaction region [Korman and Beyer, J. Acoust. Soc. Am. **84**, 339–349 (1988); **85**, 611–620 (1989)]. Theoretical results are now extended to predict the angular dependence of Doppler shift and spectral broadening when the incident primary wave components are plane wave pulses instead of continuous waves. Lighthill's equation is used to find the scattered density $\rho_s(\mathbf{x},t)$ at the sum frequency. Modeling isotropic turbulence that is stationary, homogeneous, Gaussian, and ''locally frozen,'' one obtains the ensembled average Fourier spectrum $\langle|\rho_s(\mathbf{x},\omega)|^2\rangle = B\int\int d\omega'\, d\omega'' |a_1(\omega')|^2 \times |a_2(\omega'')|^2 W(|\omega|-\omega'-\omega'')n_i n_j \Phi_{ij}(K)$, where $B=(\omega^2\rho_{01}\rho_{02}\mathbf{m}\cdot\mathbf{n}/c^3\rho_0 x)^2(2\pi cS/|\mathbf{m}|)$, ρ_{01}, ρ_{02}, and ρ_0 represent amplitudes and ambient density, $\mathbf{n}_{01}$, $\mathbf{n}_{02}$, and $\mathbf{n}$ are incident and scattered unit vectors, $\mathbf{m}=\mathbf{n}_{01}+\mathbf{n}_{02}$, $S=$ unit area, $a_{1,2}(\omega)=$ Fourier amplitude spectrum for each primary pulse, $\mathbf{K}=(|\omega|\mathbf{n}-\omega'\mathbf{n}_{01}-\omega''\mathbf{n}_{02})/c$, $W(\omega)=(1/\sqrt{2\pi}K\sigma)\times\exp[-(\omega-\mathbf{K}\cdot\mathbf{U})^2/2K^2\sigma^2]$, $\Phi_{ij}(K)$ is the turbulence energy spectrum tensor and $\mathbf{U}$, σ, and c are the mean, rms turbulent velocity, and sound speed, respectively.

Session 5pSC

Speech Communication: Second Language Learning and Use (Poster Session)

Winifred T. Strange, Chair

Ph.D. Program in Speech and Hearing, City University of New York Graduate School, 365 Fifth Avenue, New York, New York 10016

Contributed Papers

All posters will be on display from 1:30 p.m. to 4:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:30 p.m. to 2:45 p.m. and contributors of even-numbered papers will be at their posters from 2:45 p.m. to 4:00 p.m.

5pSC1. Acoustic variability and perceived foreign accent. Lesley Carmichael (Dept. of Linguist., Univ. of Washington, 1815 NE 58th St., Seattle, WA 98105, lesley@u.washington.edu)

Most people who learn a second language (L2) in late adolescence or adulthood retain some degree of foreign accent. This study investigates whether measurable intrafactor variability in the speech signal correlates with subjective ratings of degree of foreign accentedness, and whether there is an allowable deviation from native speaker norms corresponding to ratings of ''little'' or ''no'' accent. Native Korean speakers with varying degrees of English L2 proficiency and native English controls provided the stimuli. Tokens each contained one of three differences between Korean and English phonologies, potentially creating opportunities for accented speech: English contrasts between [l] and [ɹ]; between tense and lax vowels; and the presence of consonant clusters. Monolingual native English respondents indicated the degree of foreign accent they perceived. The stimuli recordings were subjected to acoustic analysis. Intrafactor variability measurements correlated with subjective perceptions of degree of foreign accent, implying that listeners are sensitive to specific, measurable variances in the acoustic signal. Consonant cluster avoidance correlated strongly with perceptions of foreign accent, followed by [l]-[ɹ] distance, and then vowel quality. Together with the measurement data, the ratings indicate that native listeners accept ranges of acoustic performance for different degrees of accentedness.

5pSC2. Variability in human judgments of foreign accent strength. Kristin Precoda (Speech Technol. and Res. Lab., SRI Intl., 333 Ravenswood Ave., Menlo Park, CA 94025, precoda@speech.sri.com)

Speech recognition-based measures of foreign accent strength have been shown to correlate well with human judgments. However, some automatic measures vary noticeably, while the variability of human judgments is generally unknown. This study examined variability in human ratings of foreign accent strength within speakers. Seventy-six Japanese adults with varying English pronunciation skills read an average of 65 sentences. Eleven native speakers of American English rated the utterances for foreign accent strength using a five-point anchored scale. Listeners were selected partly for their abilities both to use the full scale and to give consistent repeated ratings. To measure categorical variability, the entropy of each speaker's ratings by each listener was calculated. The mean equivalent number of equally likely categories was 2.6; entropy and the speaker's average rating were strongly related. This could be due to speaker characteristics (poor speakers may vary more) or to properties of the scale (the extremes may show a floor/ceiling effect). When the effect of average rating is removed, the mean equivalent number of equally likely categories falls to 1.3. Small significant correlations remain between entropy from pairs of listeners, reflecting some agreement as to which speakers vary more given their overall level of proficiency.

5pSC3. Degree of foreign accent in native Japanese speakers trained in non-native phonetic contrast discrimination. Ryan J. Conley (Dept. of Speech and Hearing Sci., Univ. of Washington, Box 357920, Seattle, WA 98195, ryanjc@u.washington.edu), Tobey L. Doeleman, Patricia K. Kuhl, and Erica B. Stevens (Univ. of Washington, Seattle, WA 98195)

Degree of foreign accent was evaluated in nine native Japanese speakers who participated in a behavioral training program to increase accuracy of perception of a non-native consonantal contrast. For all subjects, speech samples were collected after perceptual training. Additional pretraining samples were also collected from a subset of subjects. Subjects were asked to produce six sentences describing a set of visual images portraying people and objects. Images were chosen to elicit production of /r/ and /l/ sounds in different vowel context and word positions. Sentences produced by two native American English speakers were also recorded. A group of ten native American English speakers listened to digital recordings of all speakers' sentences, randomly presented via computer, and rated them for degree of foreign accent on a seven-point scale. Each Japanese speaker's scores were averaged and then correlated with their performance on discrimination of /r/ and /l/ before and after training, as well as with results from a functional magnetic resonance imaging study in which the same subjects participated [Doeleman *et al.*, J. Acoust. Soc. Am. **107**, 2803 (2000)]. These comparisons reveal how differences in degree of foreign accent correspond to behavioral and neuroimaging data. [Work supported by NICHD HD37954 to P.K.K.]

5pSC4. English phoneme and word recognition by non-native English speakers as a function of spectral resolution and English experience. Monica Padilla (Dept. of Biomed. Eng., OHE 500, Univ. of Southern California, University Park, mc-1451, Los Angeles, CA 90089-1451) and Robert V. Shannon (House Ear Inst., Los Angeles, CA 90057)

Phoneme, word, and sentence recognition was measured in listeners whose first language was either Chinese or Spanish, with varying degrees of English proficiency. Listeners were divided into four categories: (a) fully bilingual, (b) extensive English exposure after the age of 5, (c) extensive exposure between the ages of 12 and 18, and (d) extensive English exposure after the age of 18. Speech was presented at a level of 70 dB A. Listening conditions included noise (SNR of 15, 10, 5, 0 and −5 dB) and reduced spectral information (2, 4, 6, 8, and 16 frequency bands). The model of Boothroyd and Nittrouer [J. Acoust. Soc. Am. **84**, 101–114 (1998)] was applied to the data to assess the relative contributions of phonemic and lexical processing as a function of language experience. Preliminary data show that non-native listeners had more difficulty understanding words, sentences, and vowels but only a slight disadvantage for

consonants. Surprisingly, English experience had less effect on word and sentence recognition than on vowel recognition. Significantly lower performance on vowel recognition was seen even for fully bilingual listeners with reduced spectral resolution.

5pSC5. Perceptual identification training of American English /r/ and /l/ by Japanese speakers generalizes to novel stimuli and tasks. Tobey L. Doeleman (Dept. of Psych., Univ. of Washington, Seattle, WA 98195), Ryan J. Conley (Univ. of Washington, Seattle, WA 98195), John S. Pruitt (Microsoft Corp., Redmond, WA 98052), Paul Iverson (Univ. College London, London NW1 2HE, UK), Patricia K. Kuhl, and Erica B. Stevens (Univ. of Washington, Seattle, WA 98195)

This study investigated the extent to which results from /r/ and /l/ perceptual identification training generalized to identification of novel natural stimuli and discrimination of synthetic stimuli in ten native Japanese speakers. A behavioral training software program was used which incorporated factors known to affect the acquisition of phonemic distinctions, including bimodal speech cues, stimulus variability, and subject-controlled stimulus presentation with immediate feedback. Natural speech tokens were digitally manipulated to create three levels of acoustic exaggeration. These levels, and other stimulus characteristics such as the number of talkers, vowel contexts, and syllable structure, varied during each training session as a function of listener's performance. Pre- and post-training identification tasks measured generalization to natural tokens of novel talkers and vowel contexts. In addition, pre- and post-training discrimination tasks measured generalization to novel tasks and synthetic stimuli. Same–different and AXB paradigms were used to measure and compare performance on the trained non-native (/r/–/l/) contrast as well as a native (/b/–/d/) contrast using synthetic CV and VCV tokens previously identified and judged good exemplars by American English listeners. Both identification and discrimination results show generalization of learning to novel stimuli and tasks. [Work supported by NICHD HD37954 to P.K.K.]

5pSC6. Neural processes underlying perception of the English /r/–/l/ contrast by native Japanese speakers before and after identification training: An fMRI study. Keiichi Tajima (Information Sci. Div., ATR Intl., Kyoto, Japan), Daniel E. Callan, Akiko M. Callan (ATR Intl., Kyoto, Japan), Rieko Kubo, Reiko Akahane-Yamada (ATR Intl., Kyoto, Japan), Sachiko Koyama (National Inst. for Physiological Sci., Aichi, Japan), and Shinobu Masaki (ATR Intl., Kyoto, Japan)

Neural processes underlying perception of difficult and easy English phonetic contrasts by native Japanese speakers are investigated using fMRI. Although difficult (/r/–/l/, /b/–/v/) and easy (/b/–/g/) contrasts likely involve similar neural processes, it is predicted that differences exist in specific brain regions (e.g., auditory association areas) as well as in functional connectivity underlying global brain processes involved with speech. We conducted feedback-based identification training of the /r/–/l/ contrast which has been shown to improve perceptual ability by Japanese [Lively *et al*., J. Acoust. Soc. Am. **96**, 2076–2087 (1994)]. Improvements in perceptual ability are predicted to be marked by changes in the neural processes mentioned above. The /b/–/v/ contrast served as a control for which no perceptual training was given. Three experimental sessions were conducted: one for each of the contrasts. The stimuli consisted of 20 word pairs. Subjects were presented, via headphones, one stimulus every 7 s, during a 2-s silent interval between scans. A block design was employed in which five words were presented followed by five control stimuli (signal-correlated noise). The subjects' task was to press a button to identify the phoneme perceived. Both fMRI and behavioral results are discussed in relation to the aforementioned predictions.

5pSC7. Cross-linguistic perceptual assimilation of American English /r/–/l/ by Korean speakers and Japanese speakers. Ryo Komaki, Keiichi Tajima, Reiko Akahane-Yamada (Information Sci. Div., ATR Intl., Kyoto, Japan), and Youngon Choi (Duke Univ., Durham, NC)

According to the Perceptual Assimilation Model (PAM) [Best *et al.*, JEP:HPP **14**, 345–360 (1988)], the difficulties that second-language learners have in distinguishing non-native phoneme contrasts is predictable from the way learners assimilate the non-native sounds into their native phonological system. The present study examined Korean and Japanese adults' perceptual assimilation of American English (AE) /r/–/l/ contrasts in first two experiments. In experiment 1 (assimilation task), subjects listened to nonsense AE words contrasting /r/ and /l/ in five syllabic positions, and chose one of the native orthographies displayed visually. In experiment 2 (rating task), they listened to the same nonsense words, and rated the ''goodness-of-fit'' of each stimulus to the presented native orthography on the scale from 1 (least fit) to 7 (best fit). Then, in experiment 3 (identification task), the ability of Korean and Japanese to identify both real AE words and nonsense words was measured, and a significant effect of position was found in both language groups. When the assimilation and rating data from experiments 1 and 2 were interpreted in terms of the PAM categories, the predicted levels of identification accuracy were in good agreement with the identification results in experiment 3.

5pSC8. On the perception and production of Japanese stops by native speakers of Korean. Takeshi Nozawa, Mari Katayama, Midori Sasaki, and Yoshihito Ishihara (Kansai Univ. of Intl. Studies, 1-18 Aoyama Shijimi-cho, Miki-shi, Hyogo, 563-0521 Japan)

In a series of experiments, native Korean (NK) speakers' production and perception of Japanese stops was examined. It was predicted that Koreans would not produce Japanese voiced stops authentically given the fact that Korean does not contain these phonetic tokens and that this may also be due in part to their perception of these tokens [i.e., they perceive them differently than do native Japanese (NJ) speakers]. First, a production study revealed that NK produced /g/ VOT values that were significantly different than NJ subjects. Second, perceptual experiments where NK and NJ heard Japanese stops produced by NK and NJ speakers revealed that in general, the two groups were in agreement when deciding whether the tokens were voiced or voiceless. However, there was a disparity between the two in that some of the NK tokens of /g/ were perceived as /k/ by NJ but were perceived as intended /g/ tokens by NK subjects. Results of an additional perceptual experiment also demonstrated that NK subjects treat Japanese voiceless stops as equivalent to Korean aspirated stops and Japanese voiced stops as Korean plain unaspirated stops (even though the VOT values for the Japanese voiceless and Korean unaspirated plain stop are more similar).

5pSC9. Perceptual assimilation and categorical discrimination of Korean stop consonants by native Thai speakers. Ratree Wayland (Prog. in Linguist., P.O. Box 115454, 209 Yon Hall, Univ. of Florida at Gainesville, Gainesville, FL 32611-5454)

This study examined cross language perceptual assimilation and cross language categorical discrimination. Experiment 1 assessed the native Thai speakers' ability to discriminate pairs of Korean stop consonants (plain, aspirated, reinforced) at three places of articulation (bilabial, alveolar, and velar). Experiment 2 assessed the perceived relation between the three classes of Korean stop consonants and the three classes of Thai stop consonants (voiced, voiceless unaspirated, voiceless aspirated). Preliminary results suggested that, in general, native Thai speakers assimilated the Korean plain and aspirated stops to the Thai aspirated stops, and the Korean reinforced stops to the Thai unaspirated stops. Results of the categorical discrimination revealed that native Thai speakers were able to successfully discriminate pairs of Korean stop consonants and that there was no relationship between their perceptual assimilation patterns and their ability to discriminate Korean stop consonants. Since the main perceptual cue to the discrimination of Thai stops is voice onset time (VOT), these results suggest that VOT is also the main perceptual cue for the discrimination of Korean stops. Moreover, laryngealization seems to play only a minor role in the perception of Korean stop consonants.

5pSC10. Perception of English linguadental fricatives by Haitian Kreyol-African American English bilinguals. Lynda J. Felder and Winifred Strange (The City Univ. of New York, Grad. Ctr., 365 Fifth Ave., New York, NY 10016, Felder1027@aol.com)

This study examined Haitian Kreyol/English bilingual speakers' ability to discriminate non-native fricative and stop contrasts representing substitution patterns in productions of linguadental fricatives in Haitian-accented English and African American English (AAE). Theories of second language (L2) speech learning propose that much of accented speech production in an L2 is due to difficulties in perceiving speech sounds that are not contrastive in L2 learners' native language or dialect. Neither voiced nor voiceless linguadental fricatives occur in Haitian Kreyol, and Haitian L2 English learners typically substitute /t, d, or v/ for linguadental fricatives. Linguadental fricatives occur in AAE, but vary with context (linguadental, labiodental substitutions or deletions). Haitian/AAE speakers and AAE/standard AE bidialectal speakers were given a categorial discrimination test in which contrasts between linguadental fricatives and alveolar stops or labiodental fricatives in initial, intervocalic, and final position were presented in real words embedded within a short sentence. Preliminary results showed that bilingual Haitian Kreyol/AAE speakers had different patterns of perceptual errors than bidialectal speakers of AAE/SAE. Perceptual error patterns were somewhat related to substitution patterns in Haitian Kreyol and AAE. Further research is underway to investigate age of English learning and dominance of dialect effects on perception. [Work supported by NIDCD-00323.]

5pSC11. Training Mandarin and Cantonese speakers to identify English vowel contrasts: Long-term retention and effect on production. Xinchun Wang (Dept. of Linguist., Simon Fraser Univ., Burnaby, BC V5A 1S6, Canada, jxwang@sfu.ca)

Synthesized hVd vowel stimuli and naturally produced CVC minimal pairs by multiple talkers were used to train native Mandarin and Cantonese speakers to identify the English /i/–/I/, /u/–/U/, and /E/–/Q/ contrasts. In the pre- and post-tests, subjects took the identification tests on synthesized and natural stimuli, and were also recorded producing the target vowel contrasts. Results showed that subjects relied on duration cues for the /i/–/I/, contrast more consistently than they did for the other two contrasts. Training effectively shifted their attention from duration to spectral cues. Trainees' perceptual performance on natural tokens improved significantly from pretest to post-test on all three contrasts. Accuracy in generalization to new words produced by new talkers was comparable to new words by familiar talkers. The effect of perceptual learning was retained 3 months later after the training was completed. Improvement in production was observed but performance differences between pre- and post-test did not reach significance. The findings suggest that increased perceptual accuracy in identifying L2 vowel contrasts through perceptual training may not be sufficient for significant improvement in production accuracy. Future studies may look at combination of simultaneous production and perceptual training for better results.

5pSC12. Investigating a computer-based method to measure school-age children's ability to discriminate non-native speech contrasts. Jesica C. Pruitt (Dept. of Speech and Hearing Sci., Univ. of Washington, Seattle, WA 98195), John S. Pruitt (Microsoft Corp., Redmond, WA 98052), Patricia K. Kuhl, and Erica B. Stevens (Univ. of Washington, Seattle, WA 98195)

Most studies of non-native speech perception have examined adult subject populations, as tests on infants, toddlers, and older children can be methodologically difficult to conduct. Direct application of methods used for adults in the assessment of children can be problematic. Such methods can be long, mentally fatiguing, and cognitively complex. In addition to task variables, stimuli which are difficult to perceive, such as non-native speech sounds, make such tests on children all the more difficult. The present study examines a new computer-based testing method specifically designed to test the discrimination of non-native speech sounds with 5- to 15-year-old children. The design uses an oddity task that presents non-native phoneme contrasts in a way that is engaging and motivating. For baseline comparison, discrimination data were obtained from 10 AE and 10 Hindi speaking adults using four phonetically different, natural CV Hindi speech contrasts and compared to identification data previously obtained from AE and Hindi adults using the same stimuli (Pruitt, 1995) implemented with a standard adult-testing protocol. Results from pilot testing with children and adults using this new computer-based method will be discussed and compared to the baseline identification and discrimination data. [Work supported by NICHD HD37954 to P. K. Kuhl.]

5pSC13. Childhood speaking experience and perception of sentences in noise. Leah M. Knightly and Terry K. Au (Dept. of Psych., Univ. of California, Los Angeles, 1285 Franz Hall, Los Angeles, CA 90095-1563, leah@psych.ucla.edu)

Perceiving sentences in noise is a challenge for non-native speakers. Does childhood experience with a language help adult learners hear sentences in noise better? Nineteen native Spanish speakers, seven childhood speakers who had spoken Spanish fluently for several years during childhood before becoming predominantly English speaking, 10 overhearers who had overheard Spanish regularly as children but had spoken Spanish minimally, and 27 late learners who had no regular experience with Spanish until high school were asked to repeat aloud 12 Spanish sentences presented in noise (+5 dB S/N level) and then repeat the same sentences presented without noise. Performance on sentences without noise was excellent across groups, with 92%–99% of the words repeated correctly. For sentences in noise, percentages of correctly repeated words differed across groups: 88% (native speakers), 68% (childhood speakers), 49% (overhearers), and 41% (late learners) [$F(3,58)=22.417$, $p<0.001$, with sentences-without-noise performance co-varied out]. Native speakers reliably outperformed all other groups; childhood speakers reliably outperformed late learners [by HSD, $ps<0.01$]; overhearers did not differ reliably from either childhood speakers or late learners. When the test was repeated two months later, childhood speakers performed comparably to native speakers. Results suggest that childhood speaking experience may benefit speech perception when that language is later re-learned in adulthood.

5pSC14. Influence of aging on perceptual learning of second-language phonetic contrasts. Rieko Kubo, Reiko Akahane-Yamada (Information Sci. Div., ATR Intl., Kyoto, Japan), and Tsuneo Yamada (National Inst. of Multimedia Education, Chiba, Japan)

To examine the effect of aging on second-language (L2) learning, Japanese speakers of various ages, from 40s to 60s, were trained to identify American English /r/ and /l/ in a minimal pair identification task used in Lively *et al.* [J. Acoust. Soc. Am. **96**, 2076–2087 (1994)]. Pretest–post-test design was administered in which the effect of the training was evaluated by the amount of improvement in accuracy from pretest to post-test. Surprisingly, most of the subjects, even in their 60s, improved significantly from pretest to post-test. When we analyzed the data by age group (40s, 50s, and 60s), each group improved 64% to 78%, 63% to 75%, and 61% to 64%, respectively, on average. In contrast, Yamada [J. Acoust. Soc. Am. **93**, 4pSP8 (1993)] demonstrated that with the same training procedure of the same amount, Japanese adults around 20 years old improved from 60% to 85%, suggesting that the amount of improvement decreases by age. Taken together, we conclude that aging does interfere with the learning of new sounds, but not completely. Implications for the development of speech perception will be discussed.