CODEN: JASMAN 4966

THE JOURNAL of the Acoustical Society of America

Vol. 105, No. 4 April 1999

SOUNDINGS SECTION

(Continued)

CONTENTS—*Continued from preceding page*

(Continued)

CONTENTS—*Continued from preceding page*

(Continued)

CONTENTS—*Continued from preceding page*

LETTERS TO THE EDITOR

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.

Document Delivery: Copies of journal articles can be ordered from the new *Articles in Physics* online document delivery service (URL: http://www.aip.org/articles.html).

SOUNDINGS

Section Editor: Richard Stern

This front section of the *Journal* includes acoustical news, views, reviews, and general tutorial or selected research articles chosen for wide acoustical interest and written for broad acoustical readership.

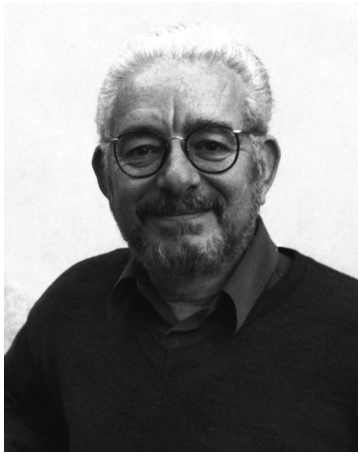
ACOUSTICAL NEWS—USA

Elaine Moran
Acoustical Society of America, 500 Sunnyside Boulevard, Woodbury, New York 11797

Editor's Note: *Readers of this Journal are asked 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.*

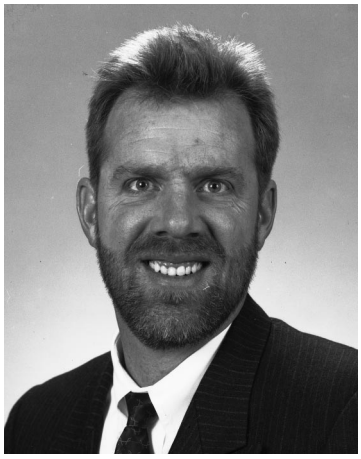
New Fellows of the Acoustical Society of America

Arlene E. Carney—For contributions to understanding speech communication in persons with cochlear implants.



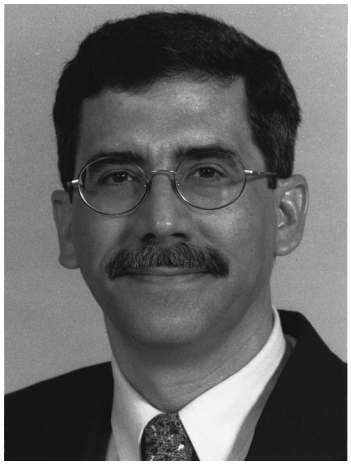
Daniel E. Commins—For contributions to architectural and environmental acoustics.

David J. Evans—For contributions to acoustical measurements and standards.



Thomas R. Howarth—For contributions to underwater acoustic transducers and systems.

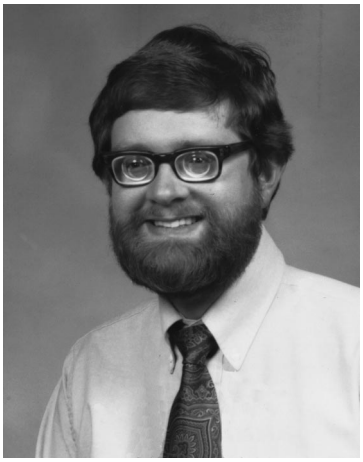
Raymond Meddis—For contributions to the modeling of auditory nerve cells and pitch perception.



Stephen T. Neely—For contributions to cochlear mechanics.

 0001-4966/99/105(4)/2057/6/$15.00 © 1999 Acoustical Society of America

Narendra T. Sanghvi—For contributions in therapeutic ultrasound.



Victor W. Sparrow—For contributions to computational acoustics and acoustics education.

1999 Acoustics Student Workshop at Tanglewood

On 29 August to 1 September 1999 a newly established Concert Hall Acoustics Summer Institute will present a Summer Student Workshop at Tanglewood, the Boston Symphony Orchestra's summer home, in Lenox, MA.

This unique educational program will be sponsored by the Newman Student Award Fund, along with a consortium of acoustical consultants and research organizations, known as the Concert Hall Research Group (CHRG), interested in encouraging a better understanding of the acoustics of music performance spaces. Sponsors include the Technical Committee of the ASA and the NCAC.

By presenting this summer program at Tanglewood, participants will have the unique opportunity to attend various concerts and to perform listening observations and acoustical measurements at the Serge Koussevitzky Concert Shed, the Seiji Ozàwa Concert Hall, and other buildings within this campus setting, under the guidance of leading concert hall designers and acoustical research engineers.

The purpose of this program is to bring students and recent graduates together with distinguished lecturers and discussion leaders in an appropriate setting, allowing participants to review and experience a wide range of facets of concert hall acoustics. The focus will be on advanced students, academic participants, practicing consultants, as well as researchers in the field of music facilities.

It is planned that at least five students will receive full-charge funded fellowships to attend the 1999 CHRG Summer Institute. Additional students, discussion leaders and other participants will be selected from applicants for a total of 30 attendees.

Applications must be submitted by 30 April 1999. Application forms and other information will be available from:

1999 Summer Institute Organizing Committee, 327F Boston Post Road, Sudbury, MA 01776, Attn. Timothy J. Foulkes, Tel: 978-443-7871, Fax: 978-443-7873, or e-mail: tfoulkes@cavtocci.com

Seven Newman Medals Awarded in 1998

In 1998, the thirteenth year of the Robert Bradford Newman Student Award Medal Program, the following students were selected to receive the medal ''For Merit in Architectural Acoustics:''

Louisa Caldas
Massachusetts Institute of Technology
Reverberation Time Optimization Tool

Max Chen
Massachusetts Institute of Technology
A Module for Calculating Noise Reduction Between Adjacent Rooms

Alton P. Chow
Syracuse University
An Urban Gesture: Defining an Identity- A Musical Performing Arts Center for the City of Taipei

Elizabeth Lee
The University of Auckland
Effects of Surface Textures of Choral Reflectors

Jason Pearson
Princeton University
Open Plan Office Acoustics: An Analysis of Theory

Joseph B. Walker
University of Florida
Sound Diffusion and the Acoustic Environment

Eugene S. Zukowski, Jr.
Roger Williams University
Reading Room in New Building for Newport Historical Society

The 1998 awards bring the total Newman Medals awarded to students to 116 at 31 institutions throughout the world since the Medal Award program began in 1985. Elizabeth Lee, the 1998 recipient, was the seventh Newman Medalist from The University of Auckland School of Architecture in New Zealand. For further information about the Newman Student Award Fund or to be added to The Newman Fund Annual Newsletter, please contact:

RBN Student Award Fund
P.O. Box 6349
Lincoln, MA 01773
Tel: 781-259-9299
Fax: 781-259-8136

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.

1999

10–12 May	AIAA/CEAS Aeroacoustics Conference, Bellevue, WA [Belur Shivashankara, The Boeing Company, P.O. Box 3707, MS 67-ML, Seattle, WA 98124-2207; Tel.: 425-234-9551; Fax: 425-237-5247; E-mail: belun.n.shivashankara@boeing.com].
4–6 June	International Hearing Aid Conference V, Iowa City, IA [Rich Tyler, Dept. Of Otolaryngology, Head and Neck Surgery, The University of Iowa, 200 Hawkins Dr., #E230 GH, Iowa City, IA 52242-1078; Tel.: 319-356-2471; Fax: 319-353-6739, E-mail: rich-tyler@uiowa.edu, WWW: http://www.medicine.uiowa.edu/otolaryngology/news.html].
6–7 June	1999 SEM Spring Conference, Cincinnati, OH [Katherine M. Ramsay, Conference Manager, Society for Experimental Mechanics, Inc., 7 School St., Bethel, CT 06801; Tel.: 203-790-6373; Fax: 203-790-4472; E-mail: meetings@sem1.com].
27–30 June	ASME Mechanics and Materials Conference, Blacksburg, VA [Mrs. Norma Guynn, Dept. of Engineering Science and Mechanics, Virginia Tech, Blacksburg, VA 24061-0219; Fax: 540-231-4574; E-mail: nguynn@vt.edu; WWW: http://www.esm.vt.edu/mmconf/].
6–11 July	1999 Clarinetfest, Ostend, Belgium [International Clarinet Association, Keith Koons, Music Dept., Univ. of Central Florida, P.O. Box 161354, Orlando, FL 32816-1354].
30 Sept.–1 Oct.	Seventh Annual Conference on the Management of the Tinnitus Patient, Iowa City, IA [Rich Tyler, Dept. of Otolaryngology, Head and Neck Surgery, The University of Iowa, 200 Hawkins Dr., #E230 GH, Iowa City, IA 52242-1078; Tel.: 319-356-2471; Fax: 319-353-6739; E-mail: rich-tyler@uiowa.edu; WWW: http://www.medicine.uiowa.edu/otolaryngology/news/news.html].
7–10 Oct.	Symposium on Occupational Hearing Loss, Philadelphia, PA [American Institute for Voice and Ear Research, Attn: Barbara-Ruth Roberts, 1721 Pine St., Philadelphia, PA 19103; Tel.: 215-545-2068; Fax: 215-735-2725]. Deadline for submission of abstracts: 1 May.
1–5 Nov.	138th meeting of the Acoustical Society of America, Columbus, OH [Acoustical Society of America, 500 Sunnyside Blvd., Woodbury, NY 11797; Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; WWW: asa.aip.org].
2–4 Dec.	ACTIVE 99, Fort Lauderdale, FL [Institute of Noise Control Engineering, P.O. Box 3206, Arlington Branch, Poughkeepsie, NY 12603; Tel.: 914-462-4006; Fax: 914-463-020; E-mail: INCEUSA@aol.com;/users.aol.com/inceusa/ince.html].
6–8 Dec.	INTER-NOISE 99, Fort Lauderdale, FL [Institute of Noise Control Engineering, P.O. Box 3206, Arlington Branch, Poughkeepsie, NY 12603; Tel.: 914-462-4006; Fax: 914-463-020; E-mail: INCEUSA@aol.com; /users.aol.com/inceusa/ince.html].

Revisions to Membership List

New Associates

Ackermann, Hermann, University of Tuebingen, Dept. of Neurology, Hoppe-Seyler-Strabe 3, Tuebingen D-72076, Germany

Alam, Sheikh K., Riverside Research Inst., Biomedical Directorate, 330 West 42nd Street, New York, NY 10036

Ambrose, Aaron M., Honda R&D Americas, Inc., ELR, 21001 State Route 739, Raymond, OH 43067-9705

Arnason, Byron T., P.O. Box 26235, Austin, TX 78755-0235

Aronov, Boris S., 22 Seabeds Way, #12, Needham, MA 02494

Asvadurov, Sergey, Schlumberger Doll Research, Old Quarry Road, Ridgefield, CT 06877-4108

Ator, Gregory A., Univ. of Kansas Medical Center, Otolaryngology, 3901 Rainbow Boulevard, Kansas City, KS 66160

Avval, Kambiz, Polyfab Corporation, Aerospace Dept., 7391 Pacific Circle, Mississauga, Ontario L5T 2A4, Canada

Basiji, David A., 4718 11th Avenue, NE, #26, Seattle, WA 98105

Benavides, Yecid A., Barrio California Calle San Francisco #80, Santa Cruz 2539, Bolivia

Benka, Stephen G., 9125 McHenry Lane, Lanham, MD 20706

Blank, Chuck, Lucas Body Systems, Central Test Lab., 5676 Industrial Park Road, Winona, MN 55987

Blomberg, Leslie D., Noise Pollution Clearinghouse, P.O. Box 1137, Montpelier, VT 05601-1137

Boe, Louis-Jean M., Institut de la Communication Parlee, Universite Stendhal, Grenoble 38040 BP 25, France

Boone, Marinus M., Voorweg 105-A, Zoetermeer 2715NG, The Netherlands

Borst, James F., CEL Instruments, 1 Westchester Drive, Milford, NH 03055

Bosmans, Ivan L., National Research Council, Inst. for Research in Construction, Montreal Road, Bldg. M-27, Ottawa, ON K1A 0R6, Canada

Bouchard, Martin, School of Information Technology and, Engineering, Univ. of Ottawa, 161 Louis Pasteur, P.O. Box 450, Station A, Ottawa, ON K1N 6N5, Canada

Briers, Rudy L. C., Zonnewende 15, Kortrijk, West Vlaanderen 8500, Belgium

Burwen, Debby L., Alaska Dept. of Fish and Game, 333 Raspberry Road, Anchorage, AK 99518

Carney, Walter L., NSWC, Crane Division, Code 7052, 300 Highway 361, Crane, IN 47522-5001

Chedeville, Pascal, 5 Rue Des Chevremonts, Nanterre 92000, France

Chen, Oiman, Shaanxi Normal University, Applied Acoustics Inst., Changan South Street, Xian, Shaanxi 710062, PR China

Cheng, Hsien K., Aerospace and Mechanical Engineering, Univ. of Southern California, University Park, Los Angeles, CA 90089-1191

Cherng, John G., University of Michigan—Dearborn, Mechanical Engineering, 4901 Evergreen Rd., Dearborn, MI 48128

Colleran, Nicholas C., 8307 Bronwood Rd., Richmond, VA 23229

Cooper, Marilyn D., Sage Technologies, 1601 N. Sepulveda Blvd., #501, Manhattan Beach, CA 90266

Cooper, Nigel P., Physiology, Univ. of Bristol, Bristol BS8 1TD U.K.

Crabtree, Brian R., DCIEM, Noise and Communications Group, Human Engrg. Sector, 1133 Sheppard Ave., West, P.O. Box 2000, Toronto, ON M3M 3B9, Canada

Cristini, Paul, CNRS, Lab. de Mechanique et D'Acoustique, 31 Chemin Joseph Aiguier, Marseille 13009, France

Cross, Lowell M., School of Music, University of Iowa, 2044 Voxman Music Bldg., Iowa City, IA 52242-1793

Dau, Torsten, Physics Dept., Univ. of Oldenburg, AG Medizinische Physik, Oldenburg 26111, Germany

DeLillo, Thomas K., Dept. of Mathematics and Statistics, Wichita, KS 67260-0033

Desarnaulds, Victor, Bureau ding. G. Monay, Avenue Vinet 25, Lausanne CH-1004, France

Devries, Diemer, Saffier 60, Berkelen Rodenrijs 2651SX, The Netherlands

Dunne, Jeff A., 12 Mark Avenue, Buckhannon, WV 26201

England, Wesley B., U.S. Army TMDE, Army Primacy Standards Lab., Acoustics and Vibration, AMSAM-TMD-SP (Bldg. 5435), Redstone Arsenal, AL 35898-5400

Falcinelli, Luca, Via Vailetta 39, Dalmine BG I-24044, Italy

Fast, Arlen J., 764 Carroll Place, Teaneck, NJ 07666

Feng, Leping, KTH, Royal Inst. of Technology, Dept. of Vehicle Engineering, Teknikringen 8, Stockholm 10044, Sweden

Fox, Warren L. J., Applied Physics Lab., University of Washington, 1013 NE 40th St., Seattle, WA 98105

Fradkin, Larissa J., School of Electrical Electronic & Information, Engineering, South Bank University, 103 Borough Road, London SE1 0AA, U.K.

Gaja, Esteban, C/Dolores Marque's 39-3, Valencia 46020, Spain

Gentry, Roger L., 22331 Mt. Ephraim Rd., Dickerson, MD 20842

Gibbs, Gary P., NASA Langley Res. Center, Structural Acoustics Branch, 2 North Dryden St., M.S. 463, Hampton, VA 23681

Gochenour, Don L., Widex, Inc., 35–53 24th Street, P.O. Box 6077, Long Island City, NY 11106-4416

Goguet, Stephan, Capital Logistic Services, 8201 Euclid Avenue, Suite 202, Manassas, VA 20111
Greaves, Robert J., Massachusetts Inst. of Technology, Earth Resources Lab., 42 Carleton St., Cambridge, MA 02142
Gudmundson, Peter, Royal Inst. of Technology, Solid Mechanics, S-10044 Stockholm, Sweden
Gummer, Anthony W., HNO-Klinik, Univ. Tubingen, Silcherstrasse 5, Tubingen 72076, Germany
Hanes, Brett E., 10369 N. Commerce Parkway, Miramar, FL 33025
Hansen, Martin, Toepholm & Westermann, Ny Westergaardsvej 25, Vaerloese DK-3500, Denmark
Hartung, Klaus, Ruhr Univ. Bochum, Inst. of Communication Acoustics, Bochum 44780, Germany
Hatangadi, Ram B., Parallel Design, Acoustics and Meterology, 365 South 52nd St., Tempe, AZ 85281-7232
Hayner, Mark A., 117 High St., Reading, MA 01867
Hengel, Peter W. J., Neurobiophysics, Univ. of Groningen, Nyenborgh 4, Groningen 9747AG, The Netherlands
Hertrich, Ingo, Cottbuser Weg 22, Rottenburg D-72108, Germany
Hirayama, Makoto J., Hewlett-Packard Laboratories Japan, 3-2-2 Sakado, Takatsu-ku, Kawasaki-shi, Kanagawa 213-0012, Japan
Hodgson, Philip N., 10155 W. Sunrise Blvd., Apt. 306, Plantation, FL 33322
Holden, Andrew P., DERA, Sensors and Processing, Winfrith Dorchester, Dorset DT2 8XJ, England
Howarth, Mark J., Building Research Establishment, Acoustic Centre, Buchnalls Lane, Garston, Watford WD2 7JR, U.K.
Hunicke, David S., 263 Bacon Point Road, Woodbury, CT 06798
Hurley, Donna C., NIST, Mail Code 853/Materials Reliability Div., 325 Broadway, Boulder, CO 80303
Imano, Kazuhiko, Electrical & Electronic Engineering, Akita University, Tegata Gakuen-cho 1-1, Akita 010-8502, Japan
Inagaki, Kenjiro, Kawasaki Heavy Industries, Ltd., Airframe Tech. Res. Dept., Gifn Technical Inst., 1 Kawasaki-cho, Kakamigahara Gifu. Pref., Japan
Ito, Akiyoshi, The Institute for Science of Labour, Research Dept., 2-8-14, Sugao, Miyamae-ku, Kawasaki Kanagawa 216-8501, Japan
Jacobs, Laurence J., Civil and Environmental Engineering, Georgia Inst. of Technology, 790 Atlantic Drive, Atlanta, GA 30332-0355
Jones, Henry E., Box 66, North, VA 23128
Juang, Biing Hwang, Acoustics and Speech Research, Bell Labs., Lucent Technologies, 2D535, 600 Mountain Avenue, Murray Hill, NJ 07974
Keiichi, Kobayashi, Okiseatec Co., Ltd., 537-5 Mito, Utiura, Numazu-shi, Shizuoka-ken 410-0223, Japan
Kelly, Jeffrey J., Dept. of Mechanical Engineering, Virginia Tech., Virginia Consortium of Engrg. and Sci., 303 Butler Farm Rd., Ste. 101, Hampton, VA 23666
Kimura, Tomonori, Mitsubishi Electric Corp., Electro-Optics & Ultrasonics Dept., Ofuna 5-1-1, Kamakura, Kanagawa 247-8501, Japan
Kirby, Raymond, Mechanical Engineering, Brunel University, Uxbridge, Middlesex UB8 3PH, England
Kirwan, Albert E., Electric Boat, Acoustics, 75 Eastern Point Rd., Groton, CT 06413
Kirwen, Jr., A. D., CCPO, Old Dominion University, Crittenton Hall, Norfolk, VA 23529
Klein, Loren L., Lucas Body Systems NA, Central Test Lab., 5676 Industrial Park Road, Winona, MN 55987
Knickrehm, Glenn A., 975 Memorial Drive, #712, Cambridge, MA 02138
Kobayashi, Teruji, 1066-34, Ichigao-cho, Aoba-ku, Yokohama-shi, Kanagawa 225-0024, Japan
Krishnaswamy, Sridhar, Mechanical Engineering, Northwestern University, 2137 Sheridan Road, Evanston, IL 60208-3020
Kumar, Arun, Indian Inst. of Technology, Delhi, Centre for Applied Res. in Electronics, Havz Khas, New Delhi 110016, India
Kundu, Tribikram, Civil Engineering & Engineering Mechanics, University of Arizona, Tucson, AZ 85721
Kushwaha, Manvir S., Inst. de Fisica, Univ. Autonoma de Puebla, Apdo. Post. J-45, Puebla 72570, Mexico
Lam, Frans-Peter A., TNO Physics and Elect. Lab., Underwater Acoustics Group, Oude Wadsdorperweg 63, P.O. Box 96864
Lawrence, Christopher F. H., 123 W. Woodglen Rd., Spartanburg, SC 29301
Lawu, Tjundewo, Faculty of Engineering, Dept. of International Development Eng., Tokyo Inst. of Technology, 2-12-1 O-okayama Meguro-ku, Tokyo 152-8552, Japan
Lei, Xun, 55 Sewall Avenue, 3A, Brookline, MA 02446
LePage, Kevin D., CMR-426, APO, AE 09613-5000
Liebenthal, Einar, Albert Einstein College of Medicine, Neuroscience, Kennedy Center 915D, 1300 Morris Park Ave., Bronx, NY 10461
Lilkendey, Robert M., Jaffe Holden Scarbrough Acoustics, Inc., 114A Washington Street, Norwalk, CT 06854
Lim, Chee W., Dept. of Mechanical Engineering, The Univ. of Hong Kong, Pokfulan Road, Hong Kong
Liu, Guanghua, Ericsson, Inc., EUS/C, 1 Mountain View Road, #1614, Lynchburg, VA 24502
Luc, Thanh T., Parsons Engineering Science, Inc., 100 West Walnut Street, Pasadena, CA 91124
Lucke, Klaus, Forschungs und Tech. Zentrum (FTZ), Oekologie Der Meeressaeuger und Voegel, Hatentoern, Buesum 25761, Germany
Macaulay, Michael C., Applied Physics Lab., University of Washington, 1013 NE 40th St., Seattle, WA 98105
Magnuson, Ronald D., 2205 Meadowbrook Drive, Austin, TX 78703
Manning, Patricia A., 108 Church Street, Winchester, MA 01890
Maruszewski, Bogdan T., Poznan Univ. of Technology, Inst. of Applied Mechanics, ul. Piotrowo 3, Poznan 60-965, Poland
Matsukawa, Mami, Deshisha Univ., Dept. of Electronics, 1–3 Tatara Miyakodani, Kyotanabe, Kyoto 610-0321, Japan
Maxon, Christopher M., Odegaard & Danneskiold Samsoe, Kroghsgade 1, Copenhagen DK-2100, Denmark
Mayer, Larry A., Ocean Mapping Group, Geodesy & Geomatics Engrg., University of New Brunswick, Fredericson, NB E3B 5A3, Canada
McLennan, Bruce R., 782 West ''M'' Street, Springfield, OR 97477
Miedema, Henk H. M., TNO Prevention and Health, Section Environment and Health, P.O. Box 2215, Leiden 2301 CE, The Netherlands
Ming, Ruisen, Dept. of Mechanical and Materials Eng., Univ. of Western Australia, Perth WA 6907, Australia
Minguez, Antonio, INSIA-UPM, Acoustics, Ctra. Valencia KM 7, Madrid 28031, Spain
Mittal, Sandeep, 7 Swamy Court, 10 Curley Street, Bangalore 560025, India
Mori, Koichi, Research Inst. of Nat. Rehab., Center for the Disabled, Dept. of Sensory and Comm. Disorders, 4-1 Namiki, Tokorozawa Saitama 359-8555, Japan
Naguib, Marc, Inst. fuer Verhaltensbiologie, Freie Univ. Berlin, Haderslebenerstr. 9, D-12163 Berlin, Germany
Nicholson, Patrick H. F., Beth Israel Deaconess Medical Ctr., Orthopaedic Biomechanics Laboratory, 330 Brookline Avenue, Rm. 115, Boston, MA 02215
Nissim, Rosalind, 61-35 98th St., Apt. 8K, Rego Park, NY 11374
Novak, Adrian A., Lucas Body Systems NA, Central Test Lab., 5676 Industrial Park Rd., P.O. Box 5649, Winona, MN 55987
Olsson, Peter, Dept. of Mechanics, Chalmers Univ. of Technology, Goteborg SE-41296, Sweden
Osler, John C., Saclant Undersea Research Centre, Viale San Bartolomeo, 400, La Spezia 19138, Italy
Oygarden, Jon, Department of Audiology, Ranheimsveien 10, Trondheim N-7005, Norway
Papadimos, Christopher A., Frank Hubach Associates, Inc., 2700 Rydin Road, Suite F, Richmond, CA 94804
Parkinson, Aaron J., University of Iowa, Otolaryngology, Head and Neck Surgery, 200 Hawkins Drive, E330 GH, Iowa City, IA 52242-1078
Pavlidou, Maria, 4 Diomidous Komninou St., Ano Pefki Attica 151-21, Greece
Pedersen, Bjarke, LIC Engineering A/S, Ehlersvej 24, Hellerup DK-2900, Denmark
Pelletier, Anik, Syprotec, Inc., 179 Brunswick Boulevard, Pointe-Claire, PQ H9R 5N2, Canada
Pena, Felix A., Communications Engineering, Inc., 8500 Cinder Bed Road, Suite 100, Newington, VA 22122
Perez, Rafael P., Univ. Politecnica De Catalunya, Fisica Aplicada, c. Jordi Girona Salgado, 1–3, Barcelona 08034, Spain
Perkins, Dwight E., Lawrence Livermore National Lab., SIPG/DTED, L-054, P.O. Box 808, Livermore, CA 94551-0808
Piske, Thorsten, Dept. of Rehabilitation Sciences, School of Health Related Professions, Univ. of Alabama at Birmingham, 503 Volker Hall, Birmingham, AL 35294-0019
Piterman, Michel, Queen's University, Dept. of Psychology, Kingston, ON K7L 3N6, Canada

Politano, Rodolfo, Rua Visconde de Ararvama no. 168, Sao Paulo 05442-080, Brazil
Popovics, John S., Civil and Architectural Engineering, 32nd and Chestnut Streets, Philadelphia, PA 19104
Psencik, Ivan, Geophysical Inst., Academy of Sciences, Seismology, Bocni II, Praha 141 31, Czech Republic
Qiao, Wenxiao, Univ. of Petroleum, China, Petroleum Resource Science, Dongying, Shandong 257062, P.R. China
Raczynski, Bohdan, 2125 Gardiner Pde., Glen Iris VIC 3146, Australia
Readhead, Mark L., DSTO, Maritime Operations Division, P.O. Box 44, Pyrmont NSW 2009, Australia
Recuero, Manuel, Inst. Univ. de Investigacion Del Automovil, Univ. Politecnica de Madrid, Carretera De Valencia KM 7, Madrid 28031, Spain
Rogers, Ian E., Nokia R&D (UK) Ltd., Speech and Audio Tag, Ashwood House, Pembroke University, Camberley, Surrey GU15 3XD, England
Santos, Jorge L. P. D., Univ. Federal de Santa Maria, Dept. De Estruturas e Construcao, Av. Roraima, S/N 0 Campus UFSM—Camobi, Santa Maria RS 97105-900, Brazil
Sapienca, Christine M., Univ. of Florida, Communication Sciences & Disorders, 63 Dauer Hall, Gainesville, FL 32611
Sarigul-Klija, Nesrin, Univ. of California at Davis, Mech. and Aero. Engineering, One Shields Ave., Davis, CA 95616-5254
Savage, Anne, Disney's Animal Kingdom, Conservation and Science, 1200 N. Savannah Circle East, Lake Buena Vista, FL 32830-1000
Scarbrough, Paul H., Jaffe Holden Scarbrough Acoustics, Inc., 114 A Washington St., Norwalk, CT 06854
Schmitt, Rainer M., Fraunhofer USA, Inc., Fraunhofer Technology Center Hialeah, 601 West 20th Street, Hialeah, FL 33010
Seale, Michael D., NASA Langley Research Center, Ultrasonics Corp., Mail Stop 231, Hampton, VA 23681-2199
Shiau, Wendy, Communicative Disorders, Speech & Hearing Ctr., College of Arts and Sciences, Univ. of Alabama, Box 870242, Tuscaloosa, AL 35487-0242
Shuman, Rebecca B., Acentech, Inc., 33 Moulton Street, Cambridge, MA 02138
Sibul, Leon H., Pennsylvania State University, ARL—Acoustics, P.O. Box 30, State College, PA 16804-0030
Simon, Henry, 3211 East Ave., Rochester, NY 14618
Skelton, Elizabeth A., 15 High Street, Meldreth, Royston, Herts SG8 6 JU, U.K.
Stangel, James E., 15233 Raintree Dr., Orland Park, IL 60462
Stepanek, Jan, Sound Studio, Music Faculty, Academy of Performing Arts Prague, Malostranske Nam. 13, Pragu 11800, Czech Republic
Sugimoto, Nobumasa, Grad. School of Engineering Sci., Osaka University, Dept. of Mech. Engrg., Toyonaka, Osaka 5608531, Japan
Sung, Shung H., General Motors Corp., Research and Development Ctr., 480-106-256, Mound Rd., 30500, Warren, MI 48090
Swanson, Douglas A., Lord Corporation, Mechanical Research and Development, 110 Lord Dr., Cary, NC 27511
Teague, Kenneth E., Arizona Music and Sound, Inc., Installed System Div., 4408 E. Speedway Blvd., Tucson, AZ 85412
Thenail, Denis, Interleuvenlaan 68–70, Research Park, Haasrode Z1, Leuven 3001, Belgium
Theunissen, Frederic E., Psychology, Univ. of California, Berkeley, 3210 Tolman, Berkeley, CA 94618
Thies, Roger M., JDB Engineering, Inc., HVAC/Acoustics, 3687 Concord Road, York, PA 17402
Truman, Christopher E., 7 Teal Close, Walkington, Beverley, East Yorkshire HU17 8TW, U.K.
Utman, Jennifer A., Center for Research in Language, Univ. of California, San Diego, 9500 Gilman Drive, Dept. 0526, La Jolla, CA 92093-0526
Van Foerster, Thomas, Springer Verlag, Physics Editorial, 175 Fifth Ave., New York, NY 10011
Ward, Darren B., Australian Defence Force Academy, School of Electrical Engrg., Northcott DVE, Canberra ACT 2600, Australia
Waterman, Elisabeth H., Poolsterstraat 24, Alphenaandenrijn NL-2402BL, The Netherlands
Wells, Andrew T., 5207 Binz-Engelman Rd., San Antonio, TX 78219-1960
Wojcik, Gregory L., Weidlinger Associates, Applied Science, 4410 El Camino Real, Suite 110, Los Altos, CA 94022
Woodger, Andrew N., Arup Acoustics, 155 Avenue of the Americas, New York, NY 10013
Yu, Man H., Aerosonics, Inc., 1601 Industrial Park Dr., P.O. Box 169, California, MO 65018
Zhang, Han, Herzog Services, Inc., 600 South Riverside Road, St. Joseph, MO 64502
Zheng, Ling, Hearing Research Center, Boston University, Biomedical Engineering, 44 Cummington Street, Boston, MA 02215
Zhong, Pei, Duke University, Mech. Engrg. and Materials Sci., 1 Science Dr., Box 90300, Durham, NC 27708
Zurk, Lisa M., Massachusetts Inst. of Technology, Lincoln Lab., 244 Wood St., Lexington, MA 02173

New Students

Adams, Barbara L., Electrical Engineering, Univ. of Alaska Fairbanks, P.O. Box 755915, Fairbanks, AK 99775
Ahkuputra, Visarut, 82/1 501 Aree-Sumphan Phaholyothin Rd., Samsennai, Phayathai, Bangkok 70400 Thailand
Amos, Nathan E., 2340 Burberry Lane, Bloomington, IN 47401
Apostolou, Antonios, Graduate Program in Acoustics, Pennsylvania State University, 217 Applied Science Building, University Park, PA 16802
Apple, Trent C., 3130 E. Villa Maria Rd., #113, Bryan, TX 77803
Barker, Brittan A., Dept. of Psychology, University of Iowa, E11 Seashore Hall, Iowa City, IA 52242-1407
Benoit Bird, Kelly J., Marine Mammal Research Program, P.O. Box 1106, Kailua, HI 96734
Bevan, Jeffrey S., 406 S. Armistad Ave., Apt. 5, Hampton, VA 23669
Boike, Kumiko T., Speech and Hearing Sciences, Univ. of Washington, 1417 NE 42nd Street, Seattle, WA 98105-6246
Bontomase, Anthony R., 430 West Foster Avenue, Apt. A, State College, PA 16801
Bouchilloux, Philippe, Magsoft Corp., 1223 Peoples Avenue, Troy, NY 12180
Brandes, T. Scott, 313 Lincoln Avenue, Apt. 99, College Station, TX 77840
Burgess, Cliff S., 225-6820 Rumble Street, Burnaby, BC V5E 4H9, Canada
Burleson, Deborah L., 3209 E. 10th St., Apt. #G-11, Bloomington, IN 47408
Carrilho Da Graca, Guilherme C., Westgate Apartment 709, 540 Memorial Drive, Cambridge, MA 02139
Chan, Jason S., 2039 Delores Street, West Covina, CA 91792
Chi, Taishih, 3416 Tulane Drive, #34, Hyattsville, MD 20783
Chiao, Kuang-Tao, 149 E. Sycamore Ave., Arcadia, CA 91006
Copeland, David B., 326998 Georgia Tech Station, Atlanta, GA 90332
De Bonte, Peter W., Worcester Polytechnic Inst., Electrical Engineering, 100 Institute Road, Worcester, MA 01609-2280
Denton, Jeannette M., Linguistics, Univ. of Chicago, 1010 East 59th Street, Chicago, IL 60637
Diaz, Jose A., 16214 Sagebrush Road, Tampa, FL 33618
Diehl, Scott R., 760 Toftrees Ave., Apt. 225, State College, PA 16803
Dilley, Laura C., Speech Communications Group, Massachusetts Inst. of Technology, 36-511, 77 Massachusetts Ave., Cambridge, MA 02139
DiMaggio, Joanee L., 22 Roe St., Melville, NY 11747
Eaton, Babette L., 609 Hilltop Drive, Bellefontaine, OH 43311
Ergin, Arif A., University of Illinois, Electrical and Computer Engrg., 1406 W. Green St., Urbana, IL 61801
Essl, Georg, Princeton University, Dept. of Computer Science, 35 Olden St., Princeton, NJ 08544
Farmer, Brenda L., University of Utah, Bioengineering, 50 South Central Campus Dr., Room 2480, Salt Lake City, UT 84112
Glosemeyer, Robin S., 1012 Emery Road, Apt. D-14, Lawrence, KS 66044
Gueorguiev, Dimitar P., Boston University, Aerospace and Mechanical, 110 Cummington St., Boston, MA 02215
Houston, Derek M., Johns Hopkins University, Psychology, 3400 North Charles/Ames Hall, Baltimore, MD 21218
Huang, Jinlan, Boston University, Aerospace and Mech. Engrg. Dept., 110 Cummington St., Room 337, Boston, MA 02215
Hurtgen, Clare M., Duke University, Physics and Music Dept., Box 99677, Durham, NC 27708-9677
Hutcheson, Florence V., 5623 Providence Rd., Virginia Beach, VA 23464
Jahromi, Omid S., Electrical and Computer Engineering, Univ. of Toronto, 10 Kings College Road, Toronto, ON M5S 3G4, Canada
Kaufman, David, EPFL, DP-IMO, Lausanne CH-1015 Switzerland
Kirk, Cecilia J., Linguistics, Univ. of Massachusetts, Amherst, South College, Amherst, MA 01003
Langlais, Michelle, 51 Rolens Drive, Apt. B5, Kingston, RI 02881
Lantz, Jennifer L., 13128 Valleywood Dr., Silver Spring, MD 20906-3959
Lanyon, Richard G., Univ. Lab. of Physiology, University of Oxford, Parks Road, Oxford OX1 3PT, U.K.
Leary, Adam P., 727 Campus View, Bloomington, IN 47408

Leibold, Lori J., 6110-4327 Lake Washington Boulevard., NE, Kirkland, WA 98033
Lemonds, David W., Hawaii Inst. of Marine Biology, Marine Mammal Res. Program, P.O. Box 1106, Kailua, HI 96734
Lister, Jennifer J., 1212 Garland Street, Mobile, AL 36618
Lohr, Kenneth R., UT Lab., Penn State Univ., 114 Hallowell Building, University Park, PA 16802
Lopes, Paulo A. C., Rue Luis de Camoes N. 9 7 DT, Lisboa 2685 Portugal
Marsh, Sarah E., Woods Hole Oceanographic Inst., 11 School Street, MS 36, Woods Hole, MA 02543
Miller, Matthew L., 6813 Kingswood Dr., Cedarburg, WI 53012
Moulton, Carey L., Mechanical and Aerospace Eng., North Carolina State Univ., Box 7910, Raleigh, NC 27695
Mourer, Gretchen A., 876 Claytor Square, Blacksburg, VA 24060
Nagao, Kyoko, Evermann Apt., Room 319, Bloomington, IN 47408
Nishi, Kanae, University of South Florida, Dept. of Psychology, 4202 East Fowler Ave., Tampa, FL 33620
O'Donovan, Jonathan J., Dept. of Electronic & Electrical Eng., Trinity College Dublin, Dublin, Ireland
Ohl, Claus-Dieter, Angerstr. 12, Gottingen D-37073, Germany
Okonak, Neil D., 918 Hamilton Ave., Latrobe, PA 15650-1601
Paeng, Dong-Guk, Pennsylvania State Univ., Acoustics, P.O. Box 30, State College, PA 16804-0030
Podesva, Robert J., Linguistics, Stanford University, Margaret Jacks Hall, Bldg. 460, Stanford, CA 94305-2150
Porter, Tyrme M., Applied Physics Lab., University of Washington, Physical/Medical Acoustics, 1013 NE 40th St., Seattle, WA 98105
Reby, David, INRA, IRGM, BP 27, Castanet 31326 France
Schwenke, Roger W., 112 Ridge Ave., State College, PA 16803-3522
Sestrade, Michel, Mathematical Physics, University College Dublin, Belfield Dublin 4, Ireland
Shi, Zhiqiang, 337713 Georgia Tech Station, Atlanta, GA 30332
Shiao, Tsung-Jieh, 17 Seventh St., Apt. #2, New Bedford, MA 02740
Sinder, Daniel J., 18A Cedar Lane, Highland Park, NJ 08904
Stankovic, Konstantine M., Otolaryngology, Harvard Medical School, 243 Charles Street, Boston, MA 02114
Stepura, Michael, Topoleva Str., 4–8, app. 102, Kiev 252049 Ukraine
Summerfield, Chris, 153 Lexington Avenue, Cambridge, MA 02138-3369
Tong, Christopher H., Dept. of Physics, Purdue University, 1396 Physics Bldg., West Lafayette, IN 47907-1396
Torre, III, Peter, Dept. of Communicative Disorders, University of Wisconsin-Madison, 1975 Willow Dr., Madison, WI 53706
Torres, Peter A., 1315 W. Lincoln Hwy., #106, DeKalb, IL 60115
Vallabha, Gautam K., 480 NW 20th Street, Apt. 205-B, Boca Raton, FL 33431
Viator, John A., Oregon Medical Laser Center, Div. of Applied Mathematics, 9205 SW Barnes Rd., Portland, OR 97225
Wage, Kathleen E., Massachusetts Inst. of Technology, Elec. and Computer Sci. EECS, Rm. 36-615, 77 Massachusetts Ave., Cambridge, MA 02139
Waters, Kendall R., Physics Dept., Washington University, Campus Box 1105, One Brookings Drive, St. Louis, MO 6313
Weber, Andrea, Max-Planck-Inst. for Psycholinguistics, Wundtlaan 1, Nijmegen 6525 XD, The Netherlands
Whitmer, William M., Parmly Hearing Institute, Loyola Univ. of Chicago, 6535 North Sheridan Road, Chicago, IL 60626
Wyatt, Sean C., Aerospace and Mech. Engrg., Boston University, 110 Cummington St., Boston, MA 02215
Yan, Xiang, Electrical & Computer Eng., Univ. of Massachusetts, Dartmouth, 285 Old Westport Road, North Dartmouth, MA 02747-2300
Yu, Alan C. L., Phonology Lab., Linguistics, Univ. of California, 1203 Dwinelle Hall, Berkeley, CA 94720
Zampolli, Mario, Boston University, Mechanical and Aerospace Engrg., 110 Cummington St., 3rd Floor, Boston, MA 02215
Zhang, Wei, Dept. of Speech Pathology, University of Washington, 1417 42nd St., NE, Seattle, WA 98105

Members Elected Fellows

A. E. Carney, C. W. Clark, D. E. Commins, A. Cowley, T. R. Howarth, R. Meddis, S. T. Neely, P. E. Rubin, N. T. Sanghvi, V. W. Sparrow, G. G. Weismer, M. Q. Wu

Associates Elected Members

C. J. Bajdek, M. D. Campbell, M. A. Clifton, P. Cobo, R. A. Day, E. Delory, M. P. De Vries, Y. T. Didenko, R. P. Dougherty, D. K. Eilar, J. J. Galvin III, S. N. Y. Gerges, J. Guo, T. A. Hamel, H-P. Herzel, D. A. K. Hewlett, P. Isberg, P. A. Johnson, R. K. Kataoka, H. Kawahara, S. A. Kostarev, M. Kumada, P. C. Laux, Y. Lu, R. G. Maev, G. A. Maksimov, M. M. Martinez, C. R. Mason, K. Nakayama, J. Pan, I. J. Russell, H. Shirai, C. W. Thorpe, D. J. Tollin, J. Vaissiere, L. P. Van Biesen, I. V. Vovk, Y. Wang

Students to Associates

M. H. Bakke, T. Bhatt, D. S. Brungart, P. T. Calamia, R. W. Chan, C-S. Chia, G. Clement, P. G. Coulter, J. A. Diaz, W. R. Drennan, J. J. Ehnert, S. A. Frivik, F. J. Gallun, J. R. Green, J. D. Harnsberger, G. J. Kenehan, J. E. Lane, A. I. Lavrentyev, C. Garbe Le Prell, D. A. Mann, S. G. Mitchell, E. Mercado III, J. Mobley, S. F. Morse, Y. Mu, I. R. Nizami, D. Norris, E. Oh, C-D. Ohl, T. A. Pitts, C. M. Richards, J. L. Rochat, S-I. Sato, C. D. Scott, R. L. Storms, K. Tajima, S. Takayanagi, C. Thodi-Petrou, S. K. Todd, P. Traykovski, D. Velea, J. M. Vieira, D. S. G. Vine, L. M. Wang, N. L. Warner, J. Zhou

Associates to Students

S. C. Conlon, A. W. Howitt

Resigned

W. R. Garner, A. C. Kibblewhite, W. Koidan—*Fellows*

A. Bedford, J. A. J. Biemond, A. S. Burgess, E. H. A. Colomb, G. Furnell, J. B. Gubelmann, I. E. Leonard, M. H. A. Lindsay, J. M. Lovrinic, A. J. Rosenheck, M. C. Schultz, W. N. Tuttle, J. Weiergang—*Members*

R. R. Andre, D. D. Bernhard, E. W. F. Druyvesteyn, S. A. Hanna, K. Hoogendoorn, J. A. Johnson, M. P. Karnell, H. J. Kim, G. S. McCall II, J. A. Posenecker, D. P. Salomon, A. Sander, W. M. Schuller, A. Sebald, M. R. Seeker, N. Sugamura, W. A. Van Donselaar, E. D. Young—*Associates*

A. A. Al-Shihail, L. Guertin, M. D. Klein—*Students*

Deceased

B. D. Cook, R. O. Fehr, P. P. Lele, D. W. Martin, A. Parvulescu, F. H. Slaymaker, V. Twersky—*Fellows*

J. C. Allred, F. T. Awbrey, W. E. Cubberly, D. M. Dickman, W. P. Etter, D. E. Goldman, S. A. Goss, J. B. Haynes, H. P. Meisinger, N. L. Mayerson, A. G. Richardson, A. Semmelink, R. M. Towne—*Members*

A. N. Borg—*Associate*

Fellows	834
Members	2865
Associates	3023
Students	948
	7670

ACOUSTICAL NEWS—INTERNATIONAL

Walter G. Mayer
Physics Department, Georgetown University, Washington, DC 20057

International Spring School on Acousto-optics joint with the Advances in Acousto-Optics Symposium, Gdańsk—May 1998

The 7th International Spring School on Acousto-optics and its Applications, joint with the 3rd Advances in Acousto-Optics Symposium took place at Jurata on the Hel Peninsula near Gdańsk, Poland. These Spring Schools are organized every three years by the Institute of Experimental Physics of the University of Gdańsk in cooperation with the Polish Acoustical Society and various other governmental and scientific organizations. The Advances in Acousto-Optic Symposia started in Paris as an annual meeting of the European Acousto-Optical Club. It has now been organized jointly with the University of Gdańsk and the Centre National Etude de Telecommunication, sponsored by the European Optical Society.

The joint meeting attracted participants from Europe, North and South America, Asia, and Africa. The main subjects of interest were physical and technical topics concerning light/sound interactions in liquids and solids, mainly in the ultrasonic range. A part of the presentations were concerned with photoacoustics and acoustic microscopy. Most of the papers dealt with the interaction of ultrasonic waves and light, in particular with phenomena of refraction, diffraction, modulation and polarization of the waves. Various applications of acousto-optic devices were discussed, such as modulators, deflectors, filters as well as topics related to integrated optics, signal processing, new materials, holographic imaging, tomography, and interferometry, all related to acousto-optics.

A plenary session of the meeting was dedicated to the memory of the late Bill D. Cook of the University of Houston who had been an active participant in the previous Spring School on Acousto-optics. The chairman of the memorial session, A. Korpel of the University of Iowa and other scientists who had cooperated with him recognized the many contributions Bill D. Cook had made to the field of acousto-optics; perhaps the most important one being the formulation of what has become known as the Klein–Cook parameter, a quantity without which it would not have been possible to evaluate a great number of experimentally observed phenomena in the field of acousto-optics. A round-table discussion was held as the last event of the meeting.

The participants had an opportunity to participate in two excursions, one to the small town of Hel, located at the tip of the peninsula, and the other to the city of Gdańsk to visit the acousto-optics laboratories at the university as well as the Old Town on the Baltic.

ANTONI S. ŚLIWIŃSKI
University of Gdańsk
80-952 Gdańsk, Poland

Vol. 1, No. 1—Noise & Health

This new quarterly publication is an inter-disciplinary international journal for all professions concerned with auditory and non-auditory noise. It is intended to cover a broad range of topics associated with noise pollution, its control, and its effects on hearing and health. The editor is D. Prasher of the University College London. Additional facts and information are available at www.vml.ucl.ac.uk/links/vmlweb/info/research/pan/index.htm

Papers published in JASJ(E)

A listing of Invited Papers and Regular Papers appearing in the latest issue of the English language version of the *Journal of the Acoustical Society of Japan*, JASJ(E), was published for the first time in the January 1995 issue of the Journal. This listing is continued below.

The January issue of JASJ(E), Vol. 20, No. 1 (1999) is a Special Issue on Control of Vehicle Pass-by Noise by Porous Asphalt Pavement. It contains seven contributions. The first two of the papers listed below are reviews.

U. Sandberg ''Low noise road surfaces—A state-of-the-art review''
S. Meiarashi ''Researches on low noise pavement in Japan''
M. Yamaguchi, H. Nakagawa, and T. Mizuno ''Sound absorption mechanism of porous asphalt pavement''
H. Ohnishi, S. Meiarashi, K. Takagi, and K. Ishikawa ''Attenuation factors of vehicle noise due to drainage asphalt pavement''
H. Hatanaka and K. Yamamoto ''Measurement and analysis of acoustic properties of drainage asphalt''
T. Iwase and R. Kawabata ''Measurement of basic acoustical properties of the porous pavement and their applications to the estimation of road traffic noise reduction''
Y. Oshino, T. Mikami, H. Ohnishi, and H. Tachibana ''Investigation into road vehicle noise reduction by drainage asphalt pavement''

Position open—New Zealand

The University of Auckland wishes to select an acoustician of international standing to fill the position of Chair in Acoustics in the School of Architecture. The vacancy arises as a result of the retirement of Professor Harold Marshall who pioneered the study of architectural acoustics in the University of Auckland. He founded the Acoustics Research Centre and the associated Acoustics Testing Service. The appointee to the Chair in Acoustics would also become Director of the Acoustics Research Center. The present focus is in the area of Building Acoustics and it is necessary that teaching duties in this aspect continue to be met, however, the position may be of interest to specialists in other parts of the discipline.

Inquiries should be sent to Michael Pritchard, Dean, Faculty of Architecture, Property and Planning, University of Auckland, Private Bag 92019, Auckland, New Zealand; Fax +64 9 373 7410.

International Meetings Calendar

Below are announcements of meetings to be held abroad. Entries preceded by an * are new or updated listings with full contact addresses given in parentheses. *Month/year* listings following other entries refer to meeting announcements, with full contact addresses, which were published in previous issues of the *Journal*.

April 1999

19–21	***Sonar Transducers 99**, Birmingham, UK. (The Institute of Acoustics, 77A St Peters' Street, St Albans, Herts. AL1 3BN, UK; Fax: +44 1727 850 553; e-mail: acoustics@clus1.ulcc.ac.uk; Web: ioa.essex.ac.uk/ioa/)
27–29	**International Conference on Vibration, Noise, and Structural Dynamics**, Venice. (Fax: +44 1785 35 35 52) *8/98*

May 1999

10–14	**4th International Conference on Theoretical and Computational Acoustics**, Trieste. (Fax: +39 40 32 70 40; Web: www.ogs.trieste.it/ictca99) *6/98*
24–26	**2nd International Conference on Emerging Technologies in NDT**, Athens. (Fax: +32 2 629 29 28; e-mail: mbourlau@vub.ac.be) *8/98*
24–27	**2nd EAA International Symposium on Hydroacoustics**, Gdańsk-Jurata. (Fax: +48 58 347 1535; Web: www.hydro.eti.pg.gda.pl) *2/99*
26–28	***National Meeting of the Italian Acoustical Association**, Genova, Italy. (Comune di Genova, Assoc. Ambiente, Via Garribaldi 9, 16100 Genova, Italy; e-mail: assambiente@comune.genova.it)
30–3	**16th International Evoked Response Audiometry Study Group Symposium**, Tromsø. (Fax: +47 77 62 73 69; e-mail: einar.laukli@rito.no) *12/98*

 0001-4966/99/105(4)/2063/2/$15.00 © 1999 Acoustical Society of America

June 1999

23–25 *__13th Rotterdam Symposium on Echocardiology__, Rotterdam, The Netherlands. (PAOG, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands; Fax: +31 10 436 7271; e-mail: secr@paog.fgg.eur.nl)

28–30 **1st International Congress of the East European Acoustical Association**, St. Petersburg. (Fax: +7 812 127 9323; e-mail: krylspb@sovam.com) *10/97*

28–1 **Joint Conference of Ultrasonics International '99 and World Congress on Ultrasonics'99** (UI99/WCU99), Lyngby. (Fax: +45 45 93 01 90; Web: www.msc.cornell.edu/~ui99) *6/98*

July 1999

4–9 **10th British Academic Conference in Otolaryngology**, London. (Fax: +44 171 404 4200) *10/97*

5–8 **6th International Congress on Sound and Vibrations**, Copenhagen. (Fax: +45 45 88 05 77; Web: www.dat.dtu.dk) *2/98*

August 1999

2–6 **International Symposium on High-Power Ultrasonics**, Vitebsk. (Fax: +375 212 24 39 53; e-mail: lpm@ita.belpak.vitebsk.by) *2/99*

September 1999

1–4 **15th International Symposium on Nonlinear Acoustics (ISNA-15)**, Göttingen. (Fax: +49 551 39 77 20; Web: www.physik3.gwdg.de/isna/) *10/97*

15–17 **British Society of Audiology Annual Conference**, Buxton. (Fax: +44 0118 935 1915; Web; www.b-s-a.demon.co.uk) *8/98*

October 1999

20–22 **Iberian Meeting of the Spanish and the Portuguese Acoustical Societies**, Avila. (Fax: +34 91 411 7651; e-mail: ssantiago@fresno.csic.es) *12/98*

March 2000

20–24 **Meeting of the German Acoustical Society (DAGA)**, Oldenburg. (Fax: +49 441 798 3698; e-mail: dega@aku.physik.uni-oldenburg.de) *10/98*

July 2000

4–7 **7th International Congress on Sound and Vibration**, Garmisch-Partenkirchen. (Fax: +49 531 295 2320; Web: www.iiav.org/icsv7.html) *12/98*

September 2000

3–6 *__5th French Congress on Acoustics—Joint Meeting of the Swiss and French Acoustical Societies__, Lausanne, Switzerland. (M.-N. Rossi, Ecole Polytechnique Fédérale, 1015 Lausanne, Switzerland; Fax: +41 21693 26 73)

October 2000

3–5 **WESTPRAC VII**, Kumamoto. (Fax: +81 96 342 3630; Web: cogni.eecs.kumamoto-u.ac.jp/others/westprac7) *6/98*

16–18 **2nd Iberoamerican Congress on Acoustics, 31st National Meeting of the Spanish Acoustical Society, and EAA Symposium**, Madrid. (Fax: +34 91 411 7651; e-mail: ssantiago@fresno.csic.es) *12/98*

16–20 **6th International Conference on Spoken Language Processing**, Beijing. (Fax: +86 10 6256 9079; e-mail: mchu@plum.ioa.ac.cn) *10/98*

September 2001

2–7 **17th International Congress on Acoustics (ICA)**, Rome. (Fax: +39 6 4424 0183; Web: www.uniroma1.it/energ/ica/html) *10/98*

BOOK REVIEWS

James F. Bartram
94 Kane Avenue, Middletown, Rhode Island 02842

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 Acoustical Oceanography

Herman Medwin and Clarence S. Clay

Academic Press, San Diego, California, 1998.
712 pp. Price $75.00.

What do shipwrecks, schools of fish, singing whales, and foaming seas have in common? They can all be detected and classified with acoustical methods and they are all described under one cover in this book. The seafloor, objects lying on or buried in the seafloor, suspended sediment, marine organisms of many kinds, eddies and turbulence, bubbles, temperature, salinity, and the sea surface all affect the manner in which sound propagates through the ocean. The purpose of this book is to describe methods that exploit the various effects so that sound can be used as a tool to infer important properties of the corresponding objects or processes in the ocean.

Why use sound? Light and other forms of electromagnetic radiation do not travel far in the ocean, giving the ocean a dark and mysterious appearance. As a result, we know, in many respects, more about the surface of the moon than of the interior of the ocean on our very own planet! Sound can travel very large distances in the ocean, especially at lower frequencies. Because of this ability, sound has been widely used as a means to probe the ocean's interior. For example, a patent on use of underwater sound was applied for shortly after the steamship Titanic sank due to its collision with an iceberg in 1912. The patent was for using sound for ''detecting the presence of large objects underwater.'' Since then, the applications using sound as a tool to study the ocean grew in number and diversity. Along with the various applications are a multitude of challenges. Both the applications and their corresponding challenges are addressed formally in this book, with many examples given.

Fundamentals of Acoustical Oceanography is written in a style appropriate for a broad audience at many levels. Much of the text is written in a simple tutorial manner so that both nonspecialists and people who are just entering the field can understand it. Furthermore, there is enough detail and references made to the literature so that the specialist can also make use of the material. The book spans areas of marine geology, marine biology, physical oceanography, and marine engineering and would be useful in applications involving ecology, commerce, and the military.

This book follows *Acoustical Oceanography*, published in 1977 by the same authors but with author order reversed. The first book was widely used and cited. It had eleven printings, was translated into Russian, and was cited routinely by scientists in varied disciplines and in many different journals. Since 1977, there have been significant advances in the area of acoustical oceanography. This new book incorporates many of the advances along with a new format. The first work reserved advanced topics for appendices at the end of the book while the present book integrates advanced material along with the rest of the text, but denotes the material as optional to read.

Acoustical oceanography, as defined by the authors early in the book, involves the so-called ''inverse problem.'' That is, given solutions to acoustic propagation, inverse methods can be applied to acoustical data collected in the ocean so that ocean parameters can be estimated. This inverse approach is in contrast to the ''forward problem'' which begins with *a priori* knowledge of the parameters. Calculations are made using these parameters to predict the manner in which sound propagates. In order to adequately perform an inverse calculation, the forward solution must be well known and the data must be properly sampled. Thus, much of the book is spent describing the forward solution, as well as various engineering aspects of the problem including sampling theory. With all such approaches under one cover, the reader is provided with substantial coverage of the important aspects involved in acoustical oceanography.

Medwin and Clay integrate diverse concepts and focus on solving realistic problems. Oceanography is described along with acoustics. Acoustical formulations are derived in terms of observable ocean parameters and there are many comparisons of theory with laboratory and ocean data. One major challenge in acoustical oceanography is due to the vast complexity of the ocean, both in the description of the ocean and of an acoustic signal traveling through it. In order to solve any acoustical problem, the boundary conditions and medium must be known. In the ocean, the spatial properties of the rough sea surface, volume heterogeneities (e.g., turbulence, bubbles, and marine organisms), and seafloor are very difficult to describe and are generally treated in a stochastic sense. Further complicating the problem is the fact that most of the properties vary over time (including certain aspects of the seafloor when the sediment is transported by ocean currents). The authors approach these difficulties by developing simplified models of the objects or processes in the ocean and identifying dominant acoustical effects in order to solve the problems.

Another challenge involves the fact that in order to solve any given problem, it is likely that, in addition to acoustics, there may be up to several areas of oceanography involved, including physical oceanography, marine biology, and marine geology. This multidisciplinary aspect of acoustical oceanography is a challenge for anyone coming from any one discipline. The authors address this issue by describing relevant fundamentals of the various branches of oceanography in the context of the corresponding acoustics formulations.

There is a broad range of approaches applied toward describing the acoustics. For example, analyses are quite frequently performed in both the time and frequency domain. Propagation in the ocean is described with both rays and modes, depending upon the conditions. Ray and modal solutions are also used in describing the acoustic scattering by volumetric objects such as fish. The descriptions of scattering by the rough boundaries, the sea surface and seafloor, use both deterministic wedge assemblages and statistical formulations to describe the interfaces. Technologies such as side-scan and multi-beam imaging sonars are discussed. In addition to fundamentals of sound propagation, various engineering aspects of transmission and reception of sound are presented including array processing, noise, filtering, and calibration. Sampling theory as applied to sampling acoustic signals and the environment is also presented. Key recent advances in acoustical oceanography are detailed, including a description of advanced multi-beam sonars for imaging of the seafloor, acoustic tomography to infer structure of ocean temperature, matched field processing to infer the location of a sound source, and multi-frequency inversions used to infer size distribution of bubbles and zooplankton.

There is a resounding theme among the diverse array of acoustical approaches and oceanographic processes or objects presented: Solving important problems in the ocean that are exceedingly complex and frequently impossible to solve with perfect accuracy may require one's willingness to cross disciplines, understand and appreciate the complexities of a naturally occurring medium, and possess the ability to combine intuition with mathematical skills.

This reviewer found the book a pleasure to read. Adding to the enjoyment was the fact that much of the book was read at the sea shore along the Atlantic coast (no phone calls that way!). The up to 20 kt winds during the

 0001-4966/99/105(4)/2065/2/$15.00 © 1999 Acoustical Society of America

reading churned up a chaotic foam-laiden sea surface along with the associated sounds of splashing. These ''data'' nicely fed the reviewer's imagination which was inspired by the authors in their insightful approaches involving the complex ocean environment.

TIMOTHY K. STANTON
Department of Applied Ocean Physics and Engineering
Woods Hole Oceanographic Institution
Woods Hole, Massachusetts 02543-1053

Acoustics in Moving Inhomogeneous Media

Vladimir E. Ostashev

E & FN Spon, London, 1997.
259 pp. Price $156.95.

The subject of acoustic propagation in a moving medium is a surprisingly complex one. Perhaps this complexity, together with the importance of the subject in many different areas of acoustics, accounts for the many alternative formulations of generalized wave equations and approximate solutions found in the literature. Making sense of these various treatments can be a perplexing ordeal. Fortunately this book provides a perfect antidote, by presenting equations and solutions for moving media derived carefully from first principles. Clarity and rigor rule throughout, as this book follows in the strong analytical tradition of other Russian physicists such as Ostashev's early mentor, V. I. Tatarskii.

The book is divided into two parts. Part I is adapted from Ostashev's original Russian text, *Sound Propagation in Moving Media*, but incorporates significant additional material. It begins with an interesting survey of the history and applications of the subject. After this initial material the emphasis of the book is primarily on analysis. A very careful derivation is provided of the basic equations for acoustical and internal gravity waves in inhomogeneous moving media. The parabolic equation for sound in a moving medium is derived, with a careful discussion of the approximations involved. Following are excellent chapters on geometrical acoustics approximations and wave theory. Validity of the concept of an ''effective'' sound speed for moving media is discussed. The chapter on wave theory provides integral representations for the field in a medium having arbitrary stratification in density, sound speed, and medium velocity; solutions for the important practical case of a point source above an impedance plane are provided. The final chapter in Part I discusses the sound field from a moving source, and the sound aberration for coordinate transformations.

Several of the sections in Part I derive from previous papers published by Ostashev and co-workers in the Soviet acoustics literature. This analytical work was not widely utilized by Western scientists, perhaps because of an increasing emphasis on numerical methods in the West. But of course numerical calculations still need to be based on correct equations, and (as noted by K. Attenborough in the Foreward) Ostashev's book is invaluable in this regard.

Part II consists of mostly new material (expanded from a single chapter in the original Russian volume) on propagation and scattering in moving random media. Like Part I, this material is authoritative and derived with commendable rigor. Some very fundamental new results are obtained for scattering cross sections and signal variations caused by random fluctuations in the medium velocity. These new results are based on the fundamental consideration that the general form of the spectrum for isotropic incompressible velocity fluctuations (such as small-scale turbulence) differs from that for isotropic scalar fluctuations. Most of the results for velocity fluctuations are obtained using equations for the scattered field taken from standard texts on propagation in random scalar fluctuations (such as Tatarskii's *The Effect of the Turbulent Atmosphere on Wave Propagation*), and then adapted to the case of velocity fluctuations. Hence this part of the book serves as a supplement to, rather than a replacement for, basic texts on scattering by random media.

The author's thorough scholarship in writing this book is particularly admirable. The many careful comparisons to the results of previous authors, along with the extensive and diverse list of references, attest to his extraordinary effort in incorporating and recognizing the work of other researchers. *Acoustics in Moving Inhomogeneous Media* is an outstanding contribution to the field of acoustics, and deserves to become a standard reference.

D. KEITH WILSON
U.S. Army Research Laboratory
2800 Powder Mill Road
Adelphi, Maryland 20783

REVIEWS OF ACOUSTICAL PATENTS

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.

Reviewers for this issue:

GEORGE L. AUGSPURGER, *Perception Incorporated, Box 39536, Los Angeles, California 90039*
ERIC E. UNGAR, *Bolt, Beranek and Newman Incorporated, 50 Moulton Street, Cambridge, Massachusetts 02238*

5,793,876

43.38.Hz METHOD FOR THE DIFFUSION OF A SOUND WITH A GIVEN DENSITY

Philippe Derogis *et al.*, assignors to France Telecom
11 August 1998 (Class 381/89); filed in France 3 July 1995

In theory, a cluster of highly directional loudspeakers could closely simulate the directivity versus frequency characteristic of any sound source. These would then be controlled by recordings of their respective gains and filter settings. The invention ''. . . can be used for acoustic installations in entertainment halls and places of sound diffusion, and also in the industrial field or in the field of sound diffusion in general.'' So much for the grand concept. One suspects that a workable implementation may be a bit trickier than simply mounting loudspeakers on a dodecahedron.—GLA

5,748,758

43.38.Ja ACOUSTIC AUDIO TRANSDUCER WITH AEROGEL DIAPHRAGM

Lawrence C. Menasco, Jr., Port Hueneme, CA and Jeffrey W. Menasco, Cardiff by the Sea, CA
5 May 1998 (Class 381/176); filed 25 January 1996

The invention is an elaborate thought experiment. Aerogels are unusual materials having extremely low densities yet significant mechanical strength. Using known techniques, aerogel can be combined with other materials to form a composite structure. Therefore, it is theoretically possible to design ''. . . an acoustic transducer that electromagnetically or electrostatically modulates an aerogel diaphragm to create a high fidelity audio speaker. The diaphragm is made of magnetic or magnetically permeable materials, or conductive materials, or a combination of both. The reciprocal arrangement can be used as a microphone or audio pickup.''—GLA

5,784,474

43.38.Ja METHOD AND CIRCUIT FOR IMPROVING THE POLAR RESPONSE OF A TWO-WAY HORN-LOADED LOUDSPEAKER SYSTEM

Paul Kohut *et al.*, assignors to Meyer Sound Laboratories, Incorporated
21 July 1998 (Class 381/97); filed 10 November 1994

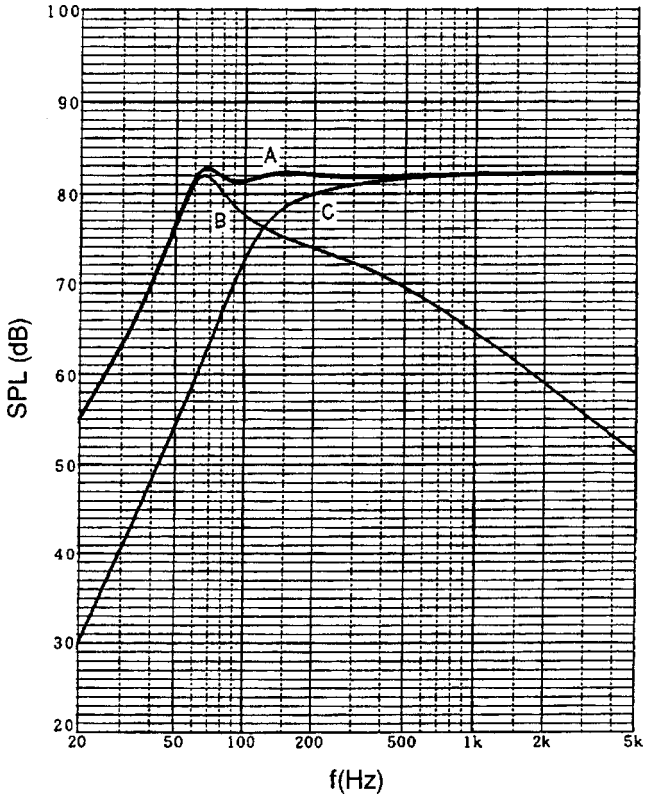
The most common loudspeaker configuration for professional playback and sound reinforcement applications consists of one or more direct radiator woofers in combination with a full-size high frequency horn. To achieve in-phase summation of sound energy on axis through the crossover region, the low frequency signal typically requires additional electronic delay, as is provided in THX theatre systems, for example. The invention is a more elaborate arrangement including amplitude and phase correction. These are adjusted to improve consistency between near-field and far-field response.—GLA

5,781,642

43.38.Ja SPEAKER SYSTEM

Shoji Tanaka and Katsuhiko Iimura, assignors to Matsushita Electric Industrial Company, Limited
14 July 1998 (Class 381/159); filed in Japan 24 April 1996

The output of an underdamped, closed box subwoofer B can be combined with that of a closed box full-range speaker C to achieve flat or nearly flat response A. So far, so good. However, the patent document asserts that



this arrangement increases efficiency by 3 dB in comparison with conventional vented systems of the same total volume. The reaction of most loudspeaker designers will be a mixture of deja vu and profound skepticism.—GLA

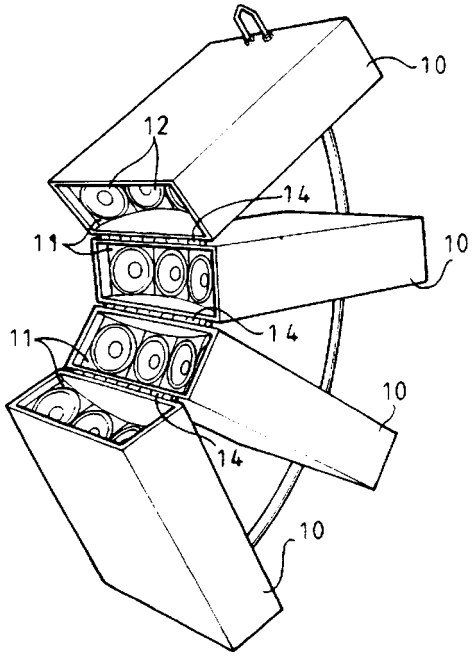
5,781,645

43.38.Ja LOUDSPEAKER SYSTEM

Christopher John Beale, assignor to SSE Hire Limited
14 July 1998 (Class 381/182); filed in the United Kingdom 28 March 1995

Loudspeaker clusters designed by C. P. Boner and others in the 1960's sometimes arranged long, mid, and short throw horns in a concave array, with the short throw horn on top. This was done for practical reasons. Dr. Boner was far too experienced to assume that ''. . . the output of the array

 0001-4966/99/105(4)/2069/6/$15.00 © 1999 Acoustical Society of America



will appear to the listener to come from a single source . . . '' or that '' . . . the comb filtering effect is substantially eliminated.'' To be fair, the commonly held notion that a *convex* array of closely spaced elements behaves as a point source is equally fallacious.—GLA

5,784,473

43.38.Ja SOUND SYSTEM

Bran Ferren, assignor to Disney Enterprises, Incorporated
21 July 1998 (Class 381/86); filed 23 February 1996

Placing small loudspeakers near a listener's head (nearphones) provides a localized sound field with a good signal-to-noise ratio, yet avoids the inconvenience of wearing headphones. The arrangement was first used by AT&T at the last New York World's Fair and is well suited to amusement park rides in general. The patent describes a seat back variant using spherical sound reflectors. ''Using this arrangement, all passengers will experience the same high quality sound, whatever the passenger's height.'' —GLA

5,793,000

43.38.Ja SPEAKER SYSTEM

Takashi Sabato *et al.*, assignors to Matsushita Electric Industrial Company, Limited
11 August 1998 (Class 181/152); filed in Japan 14 March 1995

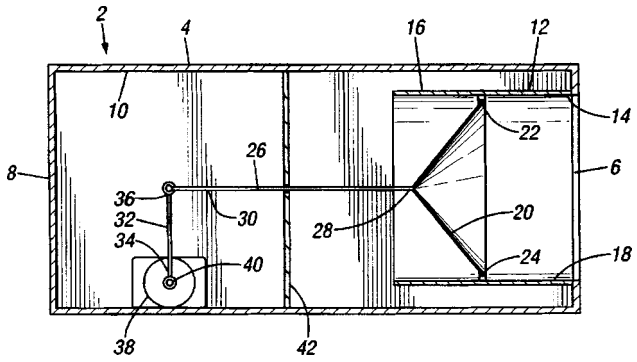
To minimize the width of a video monitor or TV set it can be advantageous to locate loudspeakers near the back of the chassis and conduct sound through waveguides to both sides of the screen. The patent describes such an arrangement including damped Helmholtz chambers to suppress unwanted organ pipe resonances.—GLA

5,802,189

43.38.Ja SUBWOOFER SPEAKER SYSTEM

Clifford L. Blodget, assignor to Samick Music Corporation
1 September 1998 (Class 381/162); filed 29 December 1995

The operational range of rotary motor **38** is restricted to less than 180 degrees. Sliding piston **20** is driven through connecting rod **26**. According to



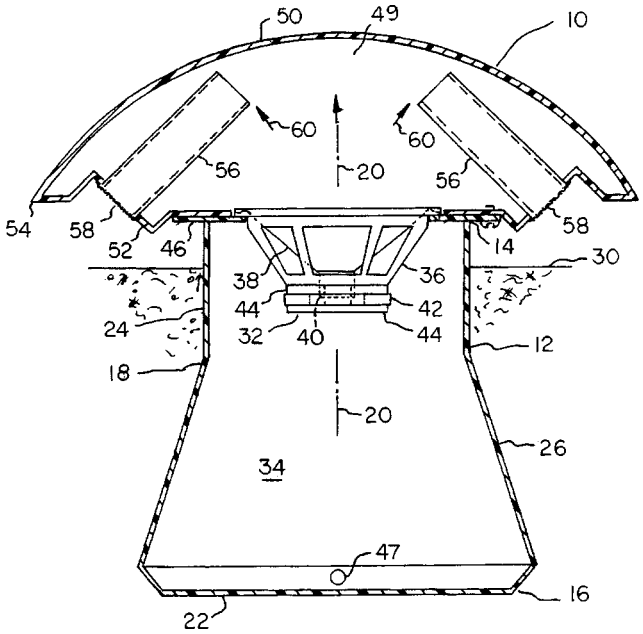
the patent document, ''A method for producing sound in response to an audio signal is also provided.'' What won't they think of next?—GLA

5,802,193

43.38.Ja OUTDOOR LOUDSPEAKER SYSTEM

William J. Kieltyka, Brunswick, ME
1 September 1998 (Class 381/188); filed 8 April 1997

The sealed back chamber **34** of this outdoor loudspeaker is mostly below ground level. A ''significant fraction'' of the sound energy in front chamber **49** is conducted to the outside world through tubes **56**. The inventor theorizes that '' . . . the action is somewhat similar to the action that takes place in a pipe organ.'' He is almost certainly right.—GLA



5,809,154

43.38.Ja PORTED LOUDSPEAKER SYSTEM AND METHOD

Matthew S. Polk, assignor to Britannia Investment Corporation
15 September 1998 (Class 381/159); filed 26 May 1995

Mr. Polk has previously patented a kind of double mushroom-shaped vent designed to minimize air turbulence. This simpler version consists of two flat disks mounted on spacers in front of and behind a simple circular port opening.—GLA

5,796,853

43.38.Lc SPEAKER AND AMPLIFIER SYSTEM

Noel Lee, assignor to Monster Cable International, Limited
18 August 1998 (Class 381/120); filed 15 December 1993

Powered loudspeaker systems have been produced commercially since the 1950's. In the past few years their popularity has increased dramatically. The idea, of course, is that a preamplifier (''voltage amplifier'') can be located remotely from an integrated power amplifier (''current amplifier'') and loudspeaker. The two are connected by lightweight cable (''said second conductors''). Well, hey, if everybody is doing it why not patent it?—GLA

5,751,822

43.38.Si AMBIENT NOISE SUPPRESSION CIRCUIT

Noboru Yamaguchi and Masao Konomi, assignors to Pan Communications, Incorporated
12 May 1998 (Class 381/94.1); filed in Japan 18 January 1994

The patent describes circuitry for use in speakerphones. ''When a microphone is located far from a mouth and is used under high ambient noise, the present invention makes it possible for the microphone to perform as if it is a microphone which is located in a quality telephone handset that cuts off ambient noise.''—GLA

5,761,314

43.38.Si AUDIO REPRODUCING APPARATUS AND HEADPHONE

Kiyofumi Inanaga and Yuji Yamada, assignors to Sony Corporation
2 June 1998 (Class 381/17); filed in Japan 27 January 1994

This long patent includes 38 pages of illustrations and describes 43 numbered inventions. The goal is more realistic stereo reproduction from headphones, including sensing and compensation for head movements. The concept of ''nearphones'' (a Bell labs designation) is explored in considerable detail, including methods of measuring and equalizing their response to provide desired transfer characteristics.—GLA

5,761,318

43.38.Si METHOD AND APPARATUS FOR MULTI-CHANNEL ACOUSTIC ECHO CANCELLATION

Suehiro Shimauchi *et al.*, assignors to Nippon Telegraph and Telephone Corporation
2 June 1998 (Class 381/66); filed in Japan 26 September 1995

Suppression of echo and feedback in multi-microphone teleconferencing systems is notoriously difficult. However, modern digital processors allow elaborate computations and corrections to be made almost instantaneously. Those who are comfortable with highly technical text and lots of heavy math will find the patent informative.—GLA

5,774,562

43.38.Si METHOD AND APPARATUS FOR DEREVERBERATION

Kenichi Furuya and Yutaka Kaneda, assignors to Nippon Telegraph and Telephone Corporation
30 June 1998 (Class 381/66); filed in Japan 25 March 1996

Low cost digital processing has prompted increasingly sophisticated and effective methods for suppressing echo and reverberation effects in speakerphones and teleconferencing systems. The invention processes the signals from two microphones in a way that does not require prior calibration. The patent is clearly written but includes heavy doses of math and transform theory.—GLA

5,790,657

43.38.Si ECHO SUPPRESSOR CAPABLE OF SUPPRESSING AN ECHO RESULTING FROM ACOUSTIC COUPLING WITHOUT SPOILING A NATURAL SOUND OF CONVERSATION

Ryuhei Fujiwara, assignor to NEC Corporation
4 August 1998 (Class 379/406); filed in Japan 26 January 1995

The invention is intended for use in a digital radio telephone system. Gain shifting is employed, based on speech detection and relative levels of send and receive signals.—GLA

5,797,124

43.38.Si VOICE-CONTROLLED VOICE MAIL HAVING RANDOM-ORDER MESSAGE RETRIEVAL BASED ON PLAYED SPOKEN IDENTIFIER LIST

James K. Walsh and Marc A. Gardner, assignors to InterVoice Limited Partnership
18 August 1998 (Class 704/275); filed 30 May 1996

This patent describes a voice-controlled voice-mail system which includes caller recognition as well as message retrieval. A caller is asked for certain identifying information, which is stored as a ''header'' along with Caller ID information (when available) and the subsequently spoken message. Previously stored headers are searched for possible similarities to the new header. When retrieving messages, the user may speak key words or phrases which have been linked to the caller's ID, providing linkages to that caller's messages.—DLR

5,748,746

43.38.Vk CEILING SPEAKER AND SIGNAL SOURCE

Yoshio Ozaki and Michael J. Kohut, assignors to Sony Corporation and Sony Cinema Products Corporation
5 May 1998 (Class 381/18); filed 27 February 1996

One might think that seven discrete channels would be enough for pretty convincing motion picture sound. But what about upper left surround and upper right surround and overhead surround? This variant of other Sony patents uses supersonic pilot tones to direct surround channels to various sets of loudspeakers.—GLA

5,761,313

43.38.Vk CIRCUIT FOR IMPROVING THE STEREO IMAGE SEPARATION OF A STEREO SIGNAL

Wayne Milton Schott, assignor to Philips Electronics North America Corporation
2 June 1998 (Class 381/1); filed 30 June 1995

Numerous cross-feed schemes have been devised to increase the subjective spread and depth of stereo reproduction from two speakers. This Philips invention employs summing, differencing, and filtering in a sophisticated, yet fairly inexpensive circuit.—GLA

5,764,777

43.38.Vk FOUR-DIMENSIONAL ACOUSTICAL AUDIO SYSTEM

Barry S. Goldfarb, assignor to BSG Laboratories, Incorporated
9 June 1998 (Class 381/27); filed 21 April 1995

This interesting configuration for enhanced stereo reproduction employs symmetrically arranged but widely separated sound sources to re-

produce various bands of frequencies (including a 4–40 kHz rear speaker).—GLA

5,784,467

43.38.Vk METHOD AND APPARATUS FOR REPRODUCING THREE-DIMENSIONAL VIRTUAL SPACE SOUND

Hiroshi Asayama, assignor to Kabushiki Kaisha Timeware
21 July 1998 (Class 381/17); filed in Japan 30 March 1995

This is a long, involved patent which will be of interest to those working with virtual reality systems. The method and apparatus are designed to obtain acoustic characteristics of full-range sound in a relatively short time even though an inexpensive computer is used. A concise description of the invention is impossible. Indeed, the patent abstract consists of a single sentence more than 100 words in length.—GLA

5,784,468

43.38.Vk SPATIAL ENHANCEMENT SPEAKER SYSTEMS AND METHODS FOR SPATIALLY ENHANCED SOUND REPRODUCTION

Arnold L. Klayman, assignor to SRS Labs, Incorporated
21 July 1998 (Class 381/24); filed 7 October 1996

A pair of conventional forward-firing loudspeakers are placed in close proximity to upward-firing speakers located immediately below. In other words, each cabinet is L-shaped, with one speaker in the vertical section and the other in the lid of the projecting base. The left forward speaker reproduces a signal comprising $(L+R)+(L-R)$ and the right forward speaker reproduces $(L+R)+(R-L)$. The left and right pedestal speakers reproduce $(L-R)$ and $(R-L)$, respectively. ''The above orientations and intensities of sounds based on the sum and difference signals provide an amazingly realistic listening experience from audio transducers positioned only at left-hand and right-hand locations.''—GLA

5,802,194

43.38.Vk STEREO LOUDSPEAKER SYSTEM WITH TWEETERS MOUNTED ON ROTATABLE ELONGATED ARMS

Makoto Yamagishi *et al.*, assignors to Sony Corporation
1 September 1998 (Class 381/188); filed in Japan 1 October 1993

A small, single package stereo system has a woofer in its main cabinet. Left and right tweeters are mounted on extended arms, like a rabbit ear antenna. This arrangement is intended to reduce the ''baffle effect'' and '' . . . improve the frequency characteristics and sound quality in a middle frequency range.'' Those interested in learning more, including the location and composition of every component and fastener, can order a copy of the patent.—GLA

5,732,802

43.40.Tm METHOD OF DAMPING VIBRATION OF STRUCTURE

Isamo Tsukagoshi, Aichi-ken, Japan
31 March 1998 (Class 188/378); filed 28 September 1996

This patent describes means for damping vibrations involving in-plane motions of grillagelike arrays of rods. Damping elements, each of which consist of a sandwich made up of two essentially rigid plates with a lawyer of viscoelastic material bonded between them, are connected between adjacent pairs of parallel rods. Each plate is mounted to each of its supporting rods via a pivot, so that it can rotate in its plane as the rods move relative to each other in their plane. The pivots on the two plates of a sandwich are at different locations along the rods, so that motion of the rods makes the two plates rotate relative to each other, causing the viscoelastic material to deform in shear and thereby to produce damping.—EEU

5,743,326

43.40.Tm METHOD OF AN APPARATUS FOR DAMPING BENDING VIBRATIONS WHILE ACHIEVING TEMPERATURE CONTROL IN BEAMS AND RELATED

Alexander H. Slocum, assignor to AESOP, Incorporated
28 April 1998 (Class 165/47); filed 5 October 1994

Although the title of this patent seems to have been truncated, perhaps because of its length, this patent does pertain to the damping of beams and structural elements whose vibrations are associated with deformations like those of beams in bending. The concept described in this patent in essence involves the use of classical viscoelastic shear damping, where shear occurs in viscoelastic elements located between an outer structural beam element and one or more internal inserts, with the neutral planes of the inserts at some distance from the neutral plane of the outer beam element. Temperature control is obtained by circulating fluids through some of the inserts or spaces not occupied by inserts.—EEU

5,756,942

43.40.Tm VIBRATION-DAMPING SECTION WITH SOUND ABSORBING MATERIAL

Toshimitsu Tanaka *et al.*, assignors to Kabushiki Kaisha Kobe Seiko Ko
26 May 1998 (Class 181/207); filed in Japan 23 October 1995

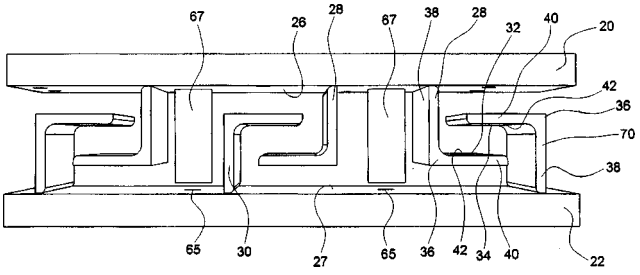
This patent pertains to panels that provide vibration damping to structures to which they are attached and acoustic absorption for spaces near these structures. These panels are intended to be attached to the undersides of steel bridges or to the walls of ship engine rooms, for example. A panel according to this patent appears somewhat like a flat box that is filled with sound absorbing material. The inside surface of the box that is to be located near the structure that is to be damped has a layer of bituminous resin fused to it. The other surface, which faces outward when the panel is installed as intended, is perforated. Bolting arrangements are provided for attaching the panel to the structure, and interlocking edge configurations are provided to connect adjacent panels to each other.—EEU

5,765,818

43.40.Tm VIBRATION-DAMPING STRUCTURE FOR USE WITH A MACHINE TOOL

David B. Sabatino, Mundelein, CA and Alan T. Hughes, South Elgin, IL
16 June 1998 (Class 267/167); filed 17 July 1996

The configuration described in this patent is intended to apply to structures that need to be both rigid and relatively highly damped. The structural configuration consists essentially of two parallel metal plates **20** and **22** with polymer concrete cast into the spaces between them. Because concrete can-



not withstand substantial tensile stresses, angles like **28** and **30** are welded to the plates before the concrete is cast in place in order to ensure that the concrete will be loaded primarily in compression.—EEU

5,769,173

43.40.Tm METHOD AND APPARATUS FOR DAMPING STRUCTURAL VIBRATIONS

J. Robert Fricke, Belmont, MA
7 July 1998 (Class 52/720.1); filed 12 June 1996

The method for damping vibration in a structural member according to this patent consists of placing granular material into intimate contact with the structure. In one type of embodiment the structural member is hollow and is filled with the granular material. In a second type of embodiment the structural member is open and an additional element is provided to hold the granular material in contact with the member. The granular materials here may be light and nonviscoelastic and may consist, for example, of polyethylene beads or glass micro-spheres. A claimed advantage, in addition to the light weight of the damping arrangement, is the flexibility one has in designing the granular material to suit specific requirements.—EEU

5,786,537

43.40.Tm VIBRATORY GROUND SURVEY SYSTEM

Nigel A. Anstey, Glen Mona, Isle of Man
28 July 1998 (Class 73/662); filed in the United Kingdom 24 July 1996

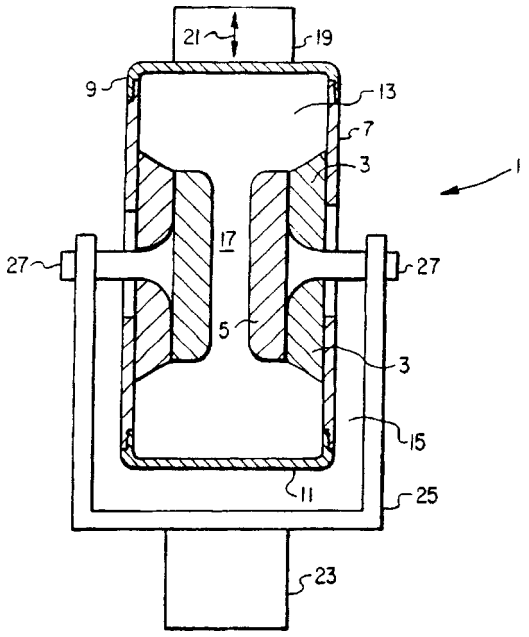
It is well known that some properties of the material at and just below the ground surface may be determined by exiting the ground surface with a quasi-sinusoidal swept-frequency signal and observing such features as resonances and anti-resonances. This patent pertains to an eccentric-mass vibrator that is brought up to speed and allowed to coast down, thus producing a frequency sweep whose rate depends on the size of the vibrator's flywheel. Since it has been observed that the vibrator decelerates more rapidly at those frequencies at which more power is radiated from the vibrator (which frequencies may be due to resonances or other factors), no separate sensing of the ground vibrations is necessary. In order to characterize the ground one essentially needs only a recording of deceleration rate versus frequency. The eccentric mass system described in this patent includes a mass that is free to move in a circular arc centered on the rotor's axis. When the rotor is accelerated, the mass is forced by its inertia to one side of the arc, where it balances a fixed mass, permitting the rotor to be brought up to speed with minimal vibration. When the rotor is free to decelerate, the inertia force moves the mass to the other end of the arc, where it serves to increase the total imbalance and thus to increase the vibratory force imparted to the ground.—EEU

5,788,029

43.40.Tm VIBRATION ISOLATION SYSTEM

Michael Reaugh Smith and Frank Bradley Stamps, assignors to Bell Helicopter
4 August 1998 (Class 188/267); filed 20 November 1996

Although this vibration isolation system was conceived for providing isolation between helicopter rotors and fuselages, it is likely to be suitable for other applications where attenuation of a single dominant frequency component is of concern. This patent is based on the Liquid Inertial Vibration Eliminator (LIVE) described in U.S. Patent No. 4,236,607 by Halwes *et al.*, a schematic cross section of which appears in the attached figure. A layer of rubber **3** acts as a liquid seal and as a spring between the inner surface of a cylinder **7** and a piston **5**. The upper and lower chambers **13** and **15** of the cylinder are connected via a passage **17** in the piston and the entire interior volume is filled with high-density, incompressible, low-viscosity fluid. The vibrating body **19** is connected at the top and the body **23** to be isolated is connected to the piston by means of a bracket **25** and lugs **27**.



Oscillations of body **19** in the direction of the arrows **21** result in relative motion between the piston **5** and cylinder **7**, causing the fluid to be pumped back and forth between chambers **13** and **15**. The inertia force corresponding to this fluid motion, which is increased by the fluid being forced through the constriction **17**, acts to oppose and possibly to cancel the externally applied force at a certain frequency. The present patent employs two LIVE configurations at right angles to each other, so as to produce isolation along two orthogonal axes. It also provides for automatic adjustment (obtained via a controller actuated by signals from an array of accelerometers) of the force cancellation frequency by magnetohydrodynamic means or by varying the dimensions of the passage **17**.—EEU

5,780,948

43.40.Tm VIBRATORY STRUCTURE, METHOD FOR CONTROLLING NATURAL FREQUENCY THEREOF AND SENSOR AND ACTUATOR ADOPTING THE VIBRATORY STRUCTURE

Ki Bang Lee, Young-ho Cho, and Ci-moo Song, assignors to Samsung Electronics Company, Limited and Korean Advanced Institute of Science and Technology
14 July 1998 (Class 310/81); filed in Republic of Korea 28 October 1995

The basic concept of this patent may be visualized by considering a mass that is attached to a support via a spring. As is well known, the system's natural frequency depends on the mass and on the spring constant. The idea here is to obtain an adjustable spring constant for adjustment of the natural frequency. An electrode is attached to the mass and a corresponding electrode to its support. When a voltage is applied between these electrodes, there results a force that varies with their relative displacement and that thus has the effect of a spring. The interaction force, and thus the spring constant, depends on the applied voltage; adjusting this voltage on may adjust the spring constant. By means of appropriate shaping of the electrodes one may obtain a linear spring characteristic—that is, linear variation of the interaction force with the relative displacement.—EEU

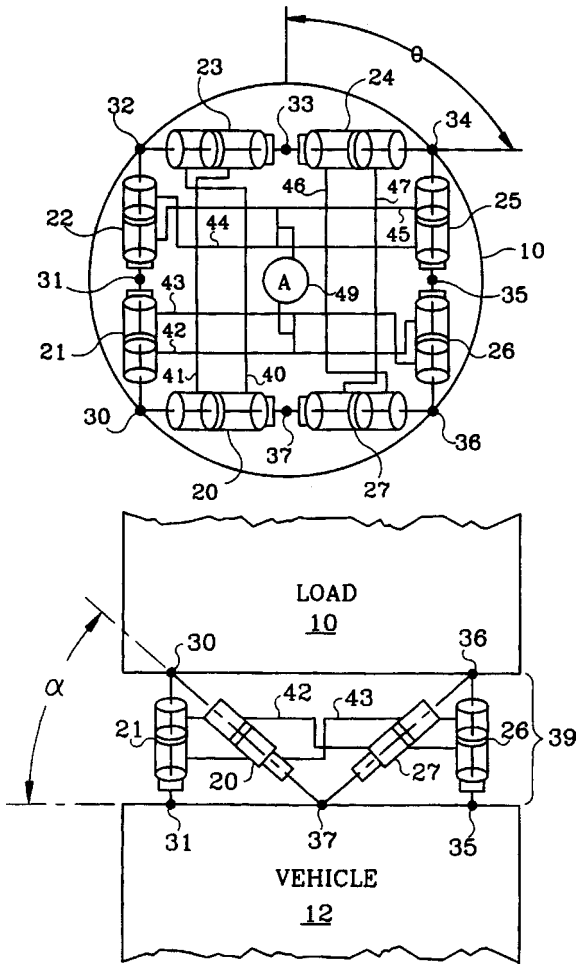
5,803,213

43.40.Tm HEAVY LOAD VIBRATION ISOLATION APPARATUS

Toren S. Davis and David A. Osterberg, assignors to Honeywell, Incorporated

8 September 1998 (Class 188/378); filed 3 February 1997

The isolation system described in this patent is intended for payloads of launch vehicles and other situations where relatively high rotational stiffness is desired in conjunction with relatively low translational stiffness. The payload **10** is isolated from the vehicle **12** via pairs of spring/damper elements **20**–**27** that are distributed around the circumference of its base, with each element inclined at an angle α and each pair connected to a common point on the payload. The ratio of the rotational to the translational stiffness can be adjusted to some degree by adjusting the angle α and the circumferential distribution of the elements. In the present patent and in a related co-pending one, increased rotational stiffness is obtained by suitable interconnection of the fluid-filled chambers of the piston/cylinder dampers as sketched in the figure that shows a circular top view of an isolator arrangement. The present patent also describes an isolation element that includes an air spring in addition to a damper, enabling different stiffnesses for supporting different loads to be obtained by adjustment of the pressure in the air springs.—EEU



5,810,131

43.40.Tm VIBRATION REDUCING METHOD

Norio Hokari *et al.*, assignors to Fuji Xerox Company, Limited

22 September 1998 (Class 188/381); filed in Japan 13 October 1995

The subject of this patent is exemplified by an electrical contact in the form of a cantilever leaf spring that touches a rotating shaft, where irregularities in the shaft's surface tend to induce vibrations in the leaf spring. Vibrations of the leaf spring are reduced by adding a secondary support near its mid-length, configured so that the support makes contact with the left spring only when that spring is deflected in one direction. The leaf spring then exhibits one set of natural frequencies when it is deflected toward the secondary support and another set when it is deflected away from that support, with the result that no resonance can occur at any excitation frequency. The concept of providing different support conditions for deflections in different directions can also be employed to reduce the vibrations of a moving medium, such as a strip of paper or film.—EEU

5,793,598

43.40.Vn MAGNETICALLY LEVITATED VIBRATION DAMPING APPARATUS

Katsuhide Watanabe *et al.*, assignors to Ebara Corporation and Kajima Corporation

11 August 1998 (Class 361/144); filed in Japan 11 August 1995

In this vibration isolation system magnetic levitation is obtained by means of an electromagnetic actuator that applies attractive forces to a flat board of a magnetic material that is fastened to the table that is to be isolated. A displacement sensor is used to detect the relative displacement between the actuator and the flat board, and an accelerometer is used to detect the absolute acceleration of the table. The signal from the displacement sensor is used via a first control rule to control the gap between the actuator and the board on which it acts. The signal from the accelerometer is used via a second control rule to control the vibration of the table, with the second control rule including stability control of the relative displacement of the table.—EEU

5,811,821

43.40.Vn SINGLE AXIS VIBRATION REDUCING SYSTEM

John D. Alexander and Michael D. Kirk, assignors to Park Scientific Instruments

22 September 1998 (Class 250/440.11); filed 9 August 1996

This patent in essence states that active single-axis translational isolation of an object may be obtained most efficiently by attaching a sensor and an actuator to the object, with the axes of the sensor and actuator parallel to the translation direction of concern and extending as nearly as possible through the center-of-mass of the object. As in any active isolation system, the actuator is activated via a suitable control system whose output depends on the sensed motion of the object. The patent claims that this sort of arrangement minimizes the coupling to rotations and to translations along other axes, and avoids the need for multiple actuators and complex control algorithms. If the motion to be controlled is the flexure (or a relative motion) of the isolated item, rather than its absolute motion, then two sensors may be used to determine the flexure of concern and the actuator is to be positioned where it has the greatest effect on this flexure.—EEU

5,769,173

43.40.Yq VIBRATION EXCITER MACHINE

Noritaka Egami *et al.*, assignors to Mitsubishi Denki Kabushiki Kaisha

23 June 1998 (Class 173/114); filed in Japan 10 June 1995

Pile driving according to this patent is accomplished by means of vibratory forces that are made to act on a pile via a pedestal that is attached to that pile. Mounted on that pedestal are magnetostrictive elements, atop which weights are fastened. Vibrations are induced by activating the magnetostrictive elements via suitable control circuits. By means of these circuits the force waveforms can also be adjusted so as to make the pile either advance into the ground or out of it.—EEU

Perfectly matched layers for elastic waves in cylindrical and spherical coordinates

Q. H. Liu[a)]
Klipsch School of Electrical and Computer Engineering, New Mexico State University, Las Cruces, New Mexico 88003

(Received 19 October 1998; revised 11 January 1999; accepted 13 January 1999)

The perfectly matched layer (PML) for elastic waves in cylindrical and spherical coordinates is developed using an improved scheme of complex coordinates. As is known for electromagnetic waves, Berenger's original PML scheme does not apply to cylindrical and spherical coordinates. The straightforward extension of the complex coordinates for elastic waves to cylindrical and spherical coordinates requires extra unknowns for time-domain solutions, wasting computer memory and computation time. The main idea of the improved scheme in this work is the use of integrated complex variables. It is shown that for three-dimensional cylindrical and spherical coordinates, this improved PML scheme requires no more unknowns than in Cartesian coordinates. The number of unknowns can be further reduced through the use of symmetry in the partial differential equations. The PML scheme allows an arbitrary inhomogeneity in the medium, and is suitable for numerical solutions of wave equations by finite-difference, finite-element, and pseudospectral methods for elastic waves in inhomogeneous media with cylindrical and spherical structures. Finite-difference time-domain (FDTD) results are shown to demonstrate the efficacy of the PML absorbing boundary condition. © *1999 Acoustical Society of America.*
[S0001-4966(99)03004-0]

PACS numbers: 43.20.Bi, 43.20.Fn, 43.20.Gp [ANN]

INTRODUCTION

Numerical solutions of partial differential equations for elastic waves require the truncation of an unbounded medium to fit into computers with a limited memory and computation time. Absorbing boundary conditions (ABCs) are needed at this truncated boundary to eliminate the reflections from this boundary to the computational domain. Considerable efforts have been made in the development of various ABCs (e.g., Refs. 1–9). However, it is still a great challenge to achieve a broadband, stable, high absorption of outgoing waves with the conventional ABCs. For example, Liao's ABC requires double precision to remain stable, yet its absorption is not ideal for cylindrical coordinates.[10]

In 1994, Berenger introduced the perfectly matched layer (PML) as a highly effective ABC for electromagnetic waves.[11] In the continuous limit, it is proven that a PML interface between a regular medium and such a fictitious perfectly matched medium completely absorbs incident waves from the regular medium regardless of its incidence angle and frequency.[11,12] This new fictitious material has since been used extensively in the finite-difference time-domain (FDTD) simulations of electromagnetic waves. Chew and Liu[13] proved that such a perfectly matched layer also exists for elastic waves in spite of the coupling of transverse (S) and longitudinal (P) waves at an elastic interface. This new ABC has been implemented for two- and three-dimensional problems. Hastings *et al.* have developed the PML ABC for two-dimensional elastic problems using potentials.[14] The PML for scalar acoustic waves has been also developed for lossy and lossless media.[15–17]

The previous efforts on PML applications in acoustic and elastic waves, however, have been concentrated on Cartesian coordinates.[13–17] For many applications, acoustic and elastic waves in inhomogeneous media with cylindrical and spherical structures are of great interest. Examples of such applications include borehole acoustic measurements[5,10,18] and global seismic modeling. Unfortunately, as demonstrated for electromagnetic waves, the original Berenger's PML scheme does not apply to cylindrical and spherical coordinates.[19–23] The straightforward applications of Berenger's PML to cylindrical coordinates will not result in a reflection-free boundary, thus the name of quasi-PML.[19,23] The idea of complex coordinates[12,13,20,21] provides a convenient way to formulate the PML for cylindrical and spherical coordinates which give a reflection-free absorbing boundary condition. Unfortunately, if applied to cylindrical and spherical coordinates in its original form, this scheme requires some extra unknowns from its counterpart in Cartesian coordinates, thus calling for more computer memory and computation time.

In this work we develop a perfectly matched layer (PML) for elastic waves in three-dimensional cylindrical and spherical coordinates without increasing the number of unknowns from the Cartesian coordinates. This formulation is based on the improved PML scheme[23] as applied to electromagnetic waves in cylindrical coordinates. It represents an improvement over the straightforward application of the method of complex coordinates. Some special 2-D elastic cases as well as the scalar acoustic problems are also discussed. This PML will be useful to frequency- and time-

a)Electronic mail: qhliu@nmsu.edu

domain numerical solutions of partial differential equations for elastic and acoustic waves.[14,15,24,10,18]

I. THE METHOD OF COMPLEX COORDINATES

Consider an inhomogeneous, anisotropic medium with mass density ρ, and the second-order elastic constant c_{jkpq} $(j,k,p,q=1,2,3)$. The linear elastic wave fields are governed by Newton's second law of motion and the generalized Hooke's law (constitutive relation). In Cartesian coordinates $(x_j, j=1,2,3)$, the equations for the particle velocity vector $\{v_j\}$ and stress tensor $\{\tau_{jk}\}$ can be written as

$$\rho\,\partial_t v_j = \partial_k \tau_{jk}, \tag{1}$$

$$\partial_t \tau_{jk} = c_{jkpq}\,\partial_q v_p, \tag{2}$$

where $\partial_k \equiv \partial/\partial x_k$, and repeated index implies summation.

Using the concept of complex coordinates in the frequency domain,[12,13,20] we introduce the complex coordinate variables

$$\tilde{x}_j = \int_0^{x_j} e_j(x_j')\,dx_j', \quad j=1,2,3, \tag{3}$$

where

$$e_j = a_j + i\omega_j/\omega, \quad a_j \geqslant 1 \ \text{ and } \ \omega_j \geqslant 0. \tag{4}$$

In the above, a time dependence of $e^{-i\omega t}$ is implied. The positive real and imaginary parts a_j and ω_j are the scaling factor and attenuation factor in the complex coordinates.[12,13,16] Using these complex coordinate variables $\tilde{x}_j$ to replace x_j in Eqs. (1) and (2), and noting that $\partial/\partial\tilde{x}_j = (1/e_j)\partial/\partial x_j$, we can obtain the frequency-domain equations in Cartesian coordinates from Eqs. (1) and (2):

$$-i\omega\rho v_j = \frac{1}{e_k}\partial_k \tau_{jk}, \tag{5}$$

$$-i\omega\tau_{jk} = c_{jkpq}\,\frac{1}{e_q}\partial_q v_p. \tag{6}$$

In the above equations, if $\omega_j > 0$, waves propagating in the complex coordinate $\tilde{x}_j$ are attenuated. Moreover, because Eqs. (5) and (6) in the complex $\{\tilde{x}_j\}$ coordinates have exactly the same forms as those in the original $\{x_j\}$ coordinates, there are no reflections even though e_j can be in general a function of x_j. Therefore, the complex coordinates can provide a perfectly matched layer to absorb outgoing waves at the edge of the computation domain in numerical solutions of these partial differential equations. In the numerical solutions, inside the region of interest we choose $e_j = 1$; while in the PML region, which acts as an absorber, we choose $e_j = a_j + i\omega_j/\omega$ with $a_j(x_j) \geqslant 1$ and $\omega_j(x_j) > 0$. In the continuous limit, there are no reflections between the regular region and the PML region.

Equations (5) and (6) can be used in frequency-domain numerical methods for elastic waves. For transient problems, we can Fourier transform Eqs. (5) and (6) into time domain. To this end, we define the split field components $v_j^{(q)}$ and $\tau_{jk}^{(q)}$ by

$$v_j = \sum_{q=1,2,3} v_j^{(q)}, \quad \tau_{jk} = \sum_{q=1,2,3} \tau_{jk}^{(q)}, \tag{7}$$

where the superscript denotes the splitting direction. Then Eqs. (5) and (6) in time domain become

$$\rho(a_q\,\partial_t + \omega_q) v_j^{(q)} = \partial_q \tau_{jq}, \quad q=1,2,3, \tag{8}$$

$$(a_q\,\partial_t + \omega_q)\tau_{jk}^{(q)} = c_{jkpq}\,\partial_q v_p, \quad q=1,2,3, \tag{9}$$

where the repeated index q does not imply summation. Noting the symmetry relation $\tau_{jk} = \tau_{kj}$, we observe that in a regular, non-PML inhomogeneous anisotropic medium $(e_j = 1)$, there are nine independent field components $\{v_j, \tau_{jk}\}$. However, the PML region contains 27 split field components (9 for $v_j^{(q)}$, and 18 for $\tau_{jk}^{(q)}$). On the other hand, if the medium is isotropic as in Ref. 13, we need only 24 split field components because $\tau_{jk}^{(q)} = 0$ for $j \neq q$ and $k \neq q$. Furthermore, if the sources for the normal stress components are identical, only three split normal stress components are necessary, making the total number of unknowns 18.

Note that the above PML formulations are derived for an inhomogeneous anisotropic medium in Cartesian coordinates. In the following sections, we will derive the PML equations for isotropic media in cylindrical and spherical coordinates.

II. CYLINDRICAL COORDINATES

A. Elastic wave equations in cylindrical coordinates

Consider an inhomogeneous isotropic medium whose Lame coefficients λ and μ, and mass density ρ are functions of space. By defining the velocity and stress vectors as

$$\mathbf{v} = [v_r, v_\theta, v_z]^T, \tag{10}$$

$$\boldsymbol{\tau} = [\tau_{rr}, \tau_{\theta\theta}, \tau_{zz}, \tau_{r\theta}, \tau_{rz}, \tau_{\theta z}]^T, \tag{11}$$

in cylindrical coordinates (r,θ,z), Newton's second law of motion and the constitutive relations can be written compactly as

$$\rho\partial_t \mathbf{v} = \left[U^{(r)} + \frac{1}{r}U^{(\theta)} + U^{(z)}\right]\boldsymbol{\tau} + \mathbf{f}, \tag{12}$$

$$\partial_t \boldsymbol{\tau} = \left[V^{(r)} + \frac{1}{r}V^{(\theta)} + V^{(z)}\right]\mathbf{v} + \mathbf{g}, \tag{13}$$

where $\mathbf{f} = [f_r, f_\theta, f_z]^T$, and $\mathbf{g} = [g_{rr}, g_{\theta\theta}, g_{zz}, g_{r\theta}, g_{rz}, g_{\theta z}]^T$ are the volume source density of body force ($\mathrm{N/m^3}$) and the volume source density of stress rate (Pa/s), respectively. In the above, the matrix operators are given by

$$U^{(r)} = \begin{bmatrix} \partial_r & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \partial_r & 0 & 0 \\ 0 & 0 & 0 & 0 & \partial_r & 0 \end{bmatrix}, \tag{14}$$

$$U^{(\theta)} = \begin{bmatrix} 1 & -1 & 0 & \partial_\theta & 0 & 0 \\ 0 & \partial_\theta & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \partial_\theta \end{bmatrix}, \tag{15}$$

$$U^{(z)}=\begin{bmatrix} 0 & 0 & 0 & 0 & \partial_z & 0 \\ 0 & 0 & 0 & 0 & 0 & \partial_z \\ 0 & 0 & \partial_z & 0 & 0 & 0 \end{bmatrix}, \tag{16}$$

$$V^{(r)}=\begin{bmatrix} (\lambda+2\mu)\partial_r & 0 & 0 \\ \lambda\partial_r & 0 & 0 \\ \lambda\partial_r & 0 & 0 \\ 0 & \mu\partial_r & 0 \\ 0 & 0 & \mu\partial_r \\ 0 & 0 & 0 \end{bmatrix}, \tag{17}$$

$$V^{(\theta)}=\begin{bmatrix} \lambda & \lambda\partial_\theta & 0 \\ (\lambda+2\mu) & (\lambda+2\mu)\partial_\theta & 0 \\ \lambda & \lambda\partial_\theta & 0 \\ \mu\partial_\theta & -\mu & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mu\partial_\theta \end{bmatrix}, \tag{18}$$

$$V^{(z)}=\begin{bmatrix} 0 & 0 & \lambda\partial_z \\ 0 & 0 & \lambda\partial_z \\ 0 & 0 & (\lambda+2\mu)\partial_z \\ 0 & 0 & 0 \\ \mu\partial_z & 0 & 0 \\ 0 & \mu\partial_z & 0 \end{bmatrix}. \tag{19}$$

Our objective is to develop a scheme based on the complex coordinates to extend the above equations so that waves in a fictitious medium with the same (λ,μ,ρ) parameters will attenuate without giving rise to reflections.

B. 3-D PML absorbing boundary condition

The PML for electromagnetic waves has been developed for cylindrical coordinates using the complex coordinates.[20–23] Here we adopt the improved complex coordinate stretching approach[23] to formulate the PML for elastic waves in cylindrical coordinates. To this end, we introduce the complex coordinate transformation

$$\tilde{r}=\int_0^r e_r(r')dr'=A_r(r)+i\Omega_r(r)/\omega, \tag{20}$$

$$\tilde{z}=\int_0^z e_z(z')dz'=A_z(z)+i\Omega_z(z)/\omega, \tag{21}$$

where $e_\eta=a_\eta+i\omega_\eta/\omega$ $(\eta=r,z)$ is the complex PML stretching variable in the frequency domain, and a_η,ω_η are real and positive functions of η. Note that unlike Cartesian coordinates which are uniform in all coordinate directions, the cylindrical and spherical coordinates are nonuniform in the radial r direction. The integrated stretching variables A_r and Ω_r compensates for this nonuniformity.

Using these complex variables to replace the original r and z variables, the frequency domain equations corresponding to Eqs. (12) and (13) can be rewritten. In particular, we do the following substitutions

$$\begin{aligned} &\partial_r\rightarrow(a_r+i\omega_r/\omega)^{-1}\partial_r, \quad r\rightarrow(A_r+i\Omega_r/\omega), \\ &\partial_z\rightarrow(a_z+i\omega_z/\omega)^{-1}\partial_z \end{aligned} \tag{22}$$

in Eqs. (12)–(19). The field components need to be split into

$$\mathbf{v}=\sum_{\eta=r,\theta,z}\mathbf{v}^{(\eta)}, \quad \boldsymbol{\tau}=\sum_{\eta=r,\theta,z}\boldsymbol{\tau}^{(\eta)}. \tag{23}$$

Then, Eqs. (12) and (13) can be written in the frequency domain (where the time-dependence of $e^{-i\omega t}$ is implied) as

$$-i\omega(\tilde{a}_\eta+i\tilde{\omega}_\eta/\omega)\rho\mathbf{v}^{(\eta)}=U^{(\eta)}\boldsymbol{\tau}, \tag{24}$$

$$-i\omega(\tilde{a}_\eta+i\tilde{\omega}_\eta/\omega)\boldsymbol{\tau}^{(\eta)}=V^{(\eta)}\mathbf{v}, \tag{25}$$

where the source terms have been omitted, and

$$\tilde{a}_\eta=\begin{cases} a_r, & \eta=r, \\ A_r, & \eta=\theta, \\ a_z, & \eta=z, \end{cases} \quad \tilde{\omega}_\eta=\begin{cases} \omega_r, & \eta=r, \\ \Omega_r, & \eta=\theta, \\ \omega_z, & \eta=z. \end{cases} \tag{26}$$

Transforming Eqs. (24) and (25) into the time domain we obtain for $\eta=(r,\theta,z)$,

$$\rho(\tilde{a}_\eta\partial_t+\tilde{\omega}_\eta)\mathbf{v}^{(\eta)}=U^{(\eta)}\boldsymbol{\tau}, \tag{27}$$

$$(\tilde{a}_\eta\partial_t+\tilde{\omega}_\eta)\boldsymbol{\tau}^{(\eta)}=V^{(\eta)}\mathbf{v}. \tag{28}$$

These equations can be solved by numerical methods such as the finite-difference method. As was recognized for electromagnetic waves,[23] with the introduction of the integrated PML variables A_r and Ω_r, the equations for cylindrical coordinates can be written in a way similar to Cartesian coordinates without introducing new split field components. Note that the difference in the split equations can all be lumped in the θ-split equations.

In Eqs. (27) and (28), the unknowns include nine velocity components $(\mathbf{v}_\zeta^{(\eta)},\eta,\zeta=r,\theta,z)$, nine normal stress components $(\boldsymbol{\tau}_{\zeta\zeta}^{(\eta)},\eta,\zeta=r,\theta,z)$, and six shear stress components $(\boldsymbol{\tau}_{\eta\zeta}^{(\eta)},\eta=r,\theta,z,\ \zeta\neq\eta)$. The other shear stress components $\boldsymbol{\tau}_{\gamma\zeta}^{(\eta)}$ for $\gamma\neq\eta$ and $\zeta\neq\eta$ are not necessary as the sixth row of Eq. (17), fifth row of Eq. (18), and fourth row of Eq. (19) are identically zero. Therefore, the total number of unknowns is 24, exactly the same as that for Cartesian coordinates. On the other hand, the straightforward extension of the complex coordinates to cylindrical coordinates requires more unknowns.[21] The introduction of integrated complex variables in Eq. (20) help reduce the number of unknowns which make this particular PML formulation attractive.

Furthermore, if the sources are applied symmetrically to all normal stress components (i.e., $g_{rr}=g_{\theta\theta}=g_{zz}$), the number of unknowns can be further reduced. This can be seen from Eqs. (17) to (19), as the split normal stress components are proportional to each other within each split direction (corresponding to the same superscript). For example, $\tau_{rr}^{(r)}$, $\tau_{\theta\theta}^{(r)}$, and $\tau_{zz}^{(r)}$ are proportional to each other, and hence only one of them is needed as an independent unknown. Therefore, the independent unknowns in Eqs. (27) and (28) are

$$\begin{aligned} &[v_r^{(r)},v_\theta^{(r)},v_z^{(r)},v_r^{(\theta)},v_\theta^{(\theta)},v_z^{(\theta)},v_r^{(z)},v_\theta^{(z)},v_z^{(z)},\tau_{rr}^{(r)}, \\ &\tau_{r\theta}^{(r)},\tau_{rz}^{(r)},\tau_{\theta\theta}^{(\theta)},\tau_{r\theta}^{(\theta)},\tau_{\theta z}^{(\theta)},\tau_{zz}^{(z)},\tau_{rz}^{(z)},\tau_{\theta z}^{(z)}]^T. \end{aligned} \tag{29}$$

There are only 18 independent unknowns and 18 equations to be solved, same as for Cartesian coordinates.[13]

C. 2-D polar coordinates

The elastic wave equations in two-dimensional polar coordinates (no z variations) can be obtained as a special case of the above formulas with $\partial_z=0$, $v_z=0$, and $\tau_{z\zeta}=0$ ($\zeta=r,\theta$). Written in their component forms, the split elastic wave equations for the 2-D polar coordinates (r,θ) are

$$\rho(a_r\,\partial_t+\omega_r)v_r^{(r)}=\partial_r\tau_{rr}, \tag{30}$$

$$\rho(a_r\,\partial_t+\omega_r)v_\theta^{(r)}=\partial_r\tau_{r\theta}, \tag{31}$$

$$\rho(A_r\,\partial_t+\Omega_r)v_r^{(\theta)}=\tau_{rr}-\tau_{\theta\theta}+\partial_\theta\tau_{r\theta}, \tag{32}$$

$$\rho(A_r\,\partial_t+\Omega_r)v_\theta^{(\theta)}=\partial_\theta\tau_{\theta\theta}+2\,\tau_{r\theta}, \tag{33}$$

$$(a_r\,\partial_t+\omega_r)\tau_{rr}^{(r)}=(\lambda+2\mu)\partial_r v_r, \tag{34}$$

$$(a_r\,\partial_t+\omega_r)\tau_{\theta\theta}^{(r)}=\lambda\,\partial_r v_r, \tag{35}$$

$$(a_r\,\partial_t+\omega_r)\tau_{zz}^{(r)}=\lambda\,\partial_r v_r, \tag{36}$$

$$(a_r\,\partial_t+\omega_r)\tau_{r\theta}^{(r)}=\mu\,\partial_r v_\theta, \tag{37}$$

$$(A_r\,\partial_t+\Omega_r)\tau_{rr}^{(\theta)}=\lambda(v_r+\partial_\theta v_\theta), \tag{38}$$

$$(A_r\,\partial_t+\Omega_r)\tau_{\theta\theta}^{(\theta)}=(\lambda+2\mu)(v_r+\partial_\theta v_\theta), \tag{39}$$

$$(A_r\,\partial_t+\Omega_r)\tau_{zz}^{(\theta)}=\lambda(v_r+\partial_\theta v_\theta), \tag{40}$$

$$(A_r\,\partial_t+\Omega_r)\tau_{r\theta}^{(\theta)}=\mu(\partial_\theta v_r-v_\theta). \tag{41}$$

Note that in Eqs. (30)–(41), only eight equations are needed for the eight independent unknowns

$$[v_r^{(r)},v_\theta^{(r)},v_r^{(\theta)},v_\theta^{(\theta)},\tau_{rr}^{(r)},\tau_{r\theta}^{(r)},\tau_{\theta\theta}^{(\theta)},\tau_{r\theta}^{(\theta)}]^T \tag{42}$$

if the sources for the normal stress components are identical. The other normal stress components can be obtained from $\tau_{rr}^{(r)}$ and $\tau_{\theta\theta}^{(\theta)}$. Therefore, there are only eight unknowns in the 2-D polar formulation.

D. 2-D axisymmetric problem

Many applications involve both axisymmetric elastic media and sources. Under such circumstances, the elastic wave equations in two-dimensional cylindrical coordinates (with no θ variations) can be obtained as a special case of the 3-D formulas with $\partial_\theta=0$, $v_\theta=0$, and $\tau_{\theta\zeta}=0$ ($\zeta=r,z$). Written in their component forms, the split elastic wave equations for the 2-D axisymmetric cylindrical coordinates (r,z) are

$$\rho(a_r\,\partial_t+\omega_r)v_r^{(r)}=\partial_r\tau_{rr}, \tag{43}$$

$$\rho(a_r\,\partial_t+\omega_r)v_z^{(r)}=\partial_r\tau_{rz}, \tag{44}$$

$$\rho(A_r\,\partial_t+\Omega_r)v_r^{(\theta)}=\tau_{rr}-\tau_{\theta\theta}, \tag{45}$$

$$\rho(A_r\,\partial_t+\Omega_r)v_z^{(\theta)}=\tau_{rz}, \tag{46}$$

$$\rho(a_z\,\partial_t+\omega_z)v_r^{(z)}=\partial_z\tau_{rz}, \tag{47}$$

$$\rho(a_z\,\partial_t+\omega_z)v_z^{(z)}=\partial_z\tau_{zz}, \tag{48}$$

$$(a_r\,\partial_t+\omega_r)\tau_{rr}^{(r)}=(\lambda+2\mu)\partial_r v_r, \tag{49}$$

$$(a_r\,\partial_t+\omega_r)\tau_{\theta\theta}^{(r)}=\lambda\,\partial_r v_r, \tag{50}$$

$$(a_r\,\partial_t+\omega_r)\tau_{zz}^{(r)}=\lambda\,\partial_r v_r, \tag{51}$$

$$(a_r\,\partial_t+\omega_r)\tau_{rz}^{(r)}=\mu\,\partial_r v_z, \tag{52}$$

$$(A_r\,\partial_t+\Omega_r)\tau_{rr}^{(\theta)}=\lambda v_r, \tag{53}$$

$$(A_r\,\partial_t+\Omega_r)\tau_{\theta\theta}^{(\theta)}=(\lambda+2\mu)v_r, \tag{54}$$

$$(A_r\,\partial_t+\Omega_r)\tau_{zz}^{(\theta)}=\lambda v_r, \tag{55}$$

$$(a_z\,\partial_t+\omega_z)\tau_{rr}^{(z)}=\lambda\,\partial_z v_z, \tag{56}$$

$$(a_z\,\partial_t+\omega_z)\tau_{\theta\theta}^{(z)}=\lambda\,\partial_z v_z, \tag{57}$$

$$(a_z\,\partial_t+\omega_z)\tau_{zz}^{(z)}=(\lambda+2\mu)\partial_z v_z, \tag{58}$$

$$(a_z\,\partial_t+\omega_z)\tau_{rz}^{(z)}=\mu\,\partial_z v_r. \tag{59}$$

Note that in Eqs. (43)–(59), only 11 equations are needed for 11 independent unknowns

$$[v_r^{(r)},v_z^{(r)},v_r^{(\theta)},v_z^{(\theta)},v_r^{(z)},v_z^{(z)},\tau_{rr}^{(r)},\tau_{rz}^{(r)},\tau_{\theta\theta}^{(\theta)},\tau_{zz}^{(z)},\tau_{rz}^{(z)}]^T \tag{60}$$

if the sources for normal stress components are identical. The other normal stress components can be obtained from $\tau_{rr}^{(r)},\tau_{\theta\theta}^{(\theta)},\tau_{zz}^{(z)}$.

III. SPHERICAL COORDINATES

A. Elastic wave equations in spherical coordinates

Consider an inhomogeneous isotropic medium whose Lame coefficients λ and μ, and mass density ρ are functions of space. In spherical coordinates (r,θ,ϕ), Newton's second law of motion and the constitutive relations can be written as

$$\rho\,\partial_t v_r=\partial_r\tau_{rr}+\frac{1}{r}\partial_\theta\tau_{r\theta}+\frac{1}{r\sin\theta}\partial_\phi\tau_{r\phi}+\frac{2\tau_{rr}-\tau_{\theta\theta}-\tau_{\phi\phi}+\tau_{r\theta}\cot\theta}{r}+f_r, \tag{61}$$

$$\rho\,\partial_t v_\theta=\frac{1}{r}\partial_\theta\tau_{\theta\theta}+\partial_r\tau_{r\theta}+\frac{1}{r\sin\theta}\partial_\phi\tau_{\theta\phi}+\frac{3\tau_{r\theta}+(\tau_{\theta\theta}-\tau_{\phi\phi})\cot\theta}{r}+f_\theta, \tag{62}$$

$$\rho\,\partial_t v_\phi=\frac{1}{r\sin\theta}\partial_\phi\tau_{\phi\phi}+\partial_r\tau_{r\phi}+\frac{1}{r}\partial_\theta\tau_{\theta\phi}+\frac{3\tau_{r\phi}+2\tau_{\theta\phi}\cot\theta}{r}+f_\phi, \tag{63}$$

$$\partial_t\tau_{rr}=\lambda\left[\partial_r v_r+\frac{2v_r}{r}+\frac{1}{r}\partial_\theta v_\theta+v_\theta\frac{\cot\theta}{r}+\frac{1}{r\sin\theta}\partial_\phi v_\phi\right]+2\mu\,\partial_r v_r+g_{rr}, \tag{64}$$

$$\partial_t\tau_{\theta\theta}=\lambda\left[\partial_r v_r+\frac{2v_r}{r}+\frac{1}{r}\partial_\theta v_\theta+v_\theta\frac{\cot\theta}{r}+\frac{1}{r\sin\theta}\partial_\phi v_\phi\right]+2\mu\left[\frac{v_r}{r}+\frac{1}{r}\partial_\theta v_\theta\right]+g_{\theta\theta}, \tag{65}$$

$$\partial_t \tau_{\phi\phi} = \lambda\left[\partial_r v_r + \frac{2v_r}{r} + \frac{1}{r}\partial_\theta v_\theta + v_\theta \frac{\cot\theta}{r} + \frac{1}{r\sin\theta}\partial_\phi v_\phi\right] + 2\mu\left[\frac{v_r}{r} + v_\theta \frac{\cot\theta}{r} + \frac{1}{r\sin\theta}\partial_\phi v_\phi\right] + g_{\phi\phi}, \tag{66}$$

$$\partial_t \tau_{r\theta} = \mu\left[\frac{1}{r}\partial_\theta v_r - \frac{v_\theta}{r} + \partial_r v_\theta\right] + g_{r\theta}, \tag{67}$$

$$\partial_t \tau_{r\phi} = \mu\left[\frac{1}{r\sin\theta}\partial_\phi v_r - \frac{v_\phi}{r} + \partial_r v_\phi\right] + g_{r\phi}, \tag{68}$$

$$\partial_t \tau_{\theta\phi} = \mu\left[\frac{1}{r\sin\theta}\partial_\phi v_\theta + \frac{1}{r}\partial_\theta v_\phi - \frac{v_\phi \cot\theta}{r}\right] + g_{\theta\phi}, \tag{69}$$

where f_α and $g_{\alpha\beta}$, $\alpha, \beta = \{r, \theta, z\}$ are the volume source density of body force (N/m^3) and the volume source density of stress rate (Pa/s), respectively.

Following the same procedures as for cylindrical coordinates and omitting the source terms, we can arrive at the following split equations for PML in spherical coordinates:

$$\rho(a_r \partial_t + \omega_r) v_r^{(r)} = \partial_r \tau_{rr}, \tag{70}$$

$$\rho(a_r \partial_t + \omega_r) v_\theta^{(r)} = \partial_r \tau_{r\theta}, \tag{71}$$

$$\rho(a_r \partial_t + \omega_r) v_\phi^{(r)} = \partial_r \tau_{r\phi}, \tag{72}$$

$$\rho(A_r \partial_t + \Omega_r) v_r^{(x)} = \partial_\theta \tau_{r\theta} + \frac{1}{\sin\theta}\partial_\phi \tau_{r\phi} + (2\tau_{rr} - \tau_{\theta\theta} - \tau_{\phi\phi} + \tau_{r\theta}\cot\theta), \tag{73}$$

$$\rho(A_r \partial_t + \Omega_r) v_\theta^{(x)} = \partial_\theta \tau_{\theta\theta} + \frac{1}{\sin\theta}\partial_\phi \tau_{\theta\phi} + [3\tau_{r\theta} + (\tau_{\theta\theta} - \tau_{\phi\phi})\cot\theta], \tag{74}$$

$$\rho(A_r \partial_t + \Omega_r) v_\phi^{(x)} = \frac{1}{\sin\theta}\partial_\phi \tau_{\phi\phi} + \partial_\theta \tau_{\theta\phi} + (3\tau_{r\phi} + 2\tau_{\theta\phi}\cot\theta), \tag{75}$$

$$(a_r \partial_t + \omega_r)\tau_{rr}^{(r)} = (\lambda + 2\mu)\partial_r v_r, \tag{76}$$

$$(a_r \partial_t + \omega_r)\tau_{\theta\theta}^{(r)} = \lambda \partial_r v_r, \tag{77}$$

$$(a_r \partial_t + \omega_r)\tau_{\phi\phi}^{(r)} = \lambda \partial_r v_r, \tag{78}$$

$$(a_r \partial_t + \omega_r)\tau_{r\theta}^{(r)} = \mu \partial_r v_\theta, \tag{79}$$

$$(a_r \partial_t + \omega_r)\tau_{r\phi}^{(r)} = \mu \partial_r v_\phi, \tag{80}$$

$$(A_r \partial_t + \Omega_r)\tau_{rr}^{(x)} = \lambda\left[2v_r + \partial_\theta v_\theta + v_\theta \cot\theta + \frac{1}{\sin\theta}\partial_\phi v_\phi\right], \tag{81}$$

$$(A_r \partial_t + \Omega_r)\tau_{\theta\theta}^{(x)} = \lambda\left[2v_r + \partial_\theta v_\theta + v_\theta \cot\theta + \frac{1}{\sin\theta}\partial_\phi v_\phi\right] + 2\mu[v_r + \partial_\theta v_\theta], \tag{82}$$

$$(A_r \partial_t + \Omega_r)\tau_{\phi\phi}^{(x)} = \lambda\left[2v_r + \partial_\theta v_\theta + v_\theta \cot\theta + \frac{1}{\sin\theta}\partial_\phi v_\phi\right] + 2\mu\left[v_r + v_\theta \cot\theta + \frac{1}{\sin\theta}\partial_\phi v_\phi\right], \tag{83}$$

$$(A_r \partial_t + \Omega_r)\tau_{r\theta}^{(x)} = \mu[\partial_\theta v_r - v_\theta], \tag{84}$$

$$(A_r \partial_t + \Omega_r)\tau_{r\phi}^{(x)} = \mu\left[\frac{1}{\sin\theta}\partial_\phi v_r - v_\phi\right], \tag{85}$$

$$(A_r \partial_t + \Omega_r)\tau_{\theta\phi}^{(x)} = \mu\left[\frac{1}{\sin\theta}\partial_\phi v_\theta + \partial_\theta v_\phi - v_\phi \cot\theta\right], \tag{86}$$

where the superscript (x) denotes the splitting in (θ,ϕ) directions. There are 17 unknowns in Eqs. (70)–(86). However, among these split field components, only the following 15 independent unknowns

$$[v_r^{(r)}, v_\theta^{(r)}, v_\phi^{(r)}, v_r^{(x)}, v_\theta^{(x)}, v_\phi^{(x)}, \tau_{rr}^{(r)}, \tau_{r\theta}^{(r)}, \tau_{r\phi}^{(r)}, \tau_{rr}^{(x)}, \tau_{\theta\theta}^{(x)}, \tau_{\phi\phi}^{(x)}, \tau_{r\theta}^{(x)}, \tau_{r\phi}^{(x)}, \tau_{\theta\phi}^{(x)}]^T \tag{87}$$

are needed if $g_{rr} = g_{\theta\theta} = g_{\phi\phi}$. It is obvious that with this formulation, the number of unknowns for spherical coordinates is smaller than the cylindrical coordinates as only one complex coordinate variable ($\tilde{r}$) is needed in spherical coordinates.

A further reduction in the number of unknowns is possible if one replaces the split equations (81)–(83) by the following equations:

$$(A_r \partial_t + \Omega_r)\tau_{rr}^{(\theta)} = \lambda[v_r + \partial_\theta v_\theta], \tag{88}$$

$$(A_r \partial_t + \Omega_r)\tau_{\theta\theta}^{(\theta)} = (\lambda + 2\mu)[v_r + \partial_\theta v_\theta], \tag{89}$$

$$(A_r \partial_t + \Omega_r)\tau_{\phi\phi}^{(\theta)} = \lambda[v_r + \partial_\theta v_\theta], \tag{90}$$

$$(A_r \partial_t + \Omega_r)\tau_{rr}^{(\phi)} = \lambda\left[v_r + v_\theta \cot\theta + \frac{1}{\sin\theta}\partial_\phi v_\phi\right], \tag{91}$$

$$(A_r \partial_t + \Omega_r)\tau_{\theta\theta}^{(\phi)} = \lambda\left[v_r + v_\theta \cot\theta + \frac{1}{\sin\theta}\partial_\phi v_\phi\right], \tag{92}$$

$$(A_r \partial_t + \Omega_r)\tau_{\phi\phi}^{(\phi)} = (\lambda + 2\mu)\left[v_r + v_\theta \cot\theta + \frac{1}{\sin\theta}\partial_\phi v_\phi\right]. \tag{93}$$

It is obvious that $\tau_{rr}^{(\theta)}$, $\tau_{\theta\theta}^{(\theta)}$, and $\tau_{\phi\phi}^{(\theta)}$ are proportional to each other, as are $\tau_{rr}^{(\phi)}$, $\tau_{\theta\theta}^{(\phi)}$, and $\tau_{\phi\phi}^{(\phi)}$. Therefore, only 14 unknowns and equations are needed for

$$[v_r^{(r)}, v_\theta^{(r)}, v_\phi^{(r)}, v_r^{(x)}, v_\theta^{(x)}, v_\phi^{(x)}, \tau_{rr}^{(r)}, \tau_{r\theta}^{(r)}, \tau_{r\phi}^{(r)}, \tau_{\theta\theta}^{(\theta)}, \tau_{\phi\phi}^{(\phi)}, \tau_{r\theta}^{(x)}, \tau_{r\phi}^{(x)}, \tau_{\theta\phi}^{(x)}]^T. \tag{94}$$

Thus the number of unknowns is smaller than for Cartesian coordinates.

IV. SCALAR ACOUSTIC WAVES

We can regard the scalar acoustic waves as a special case of the elastic waves when $c_{jklm}=\lambda\delta_{jk}\delta_{lm}$ and $\tau_{jk}=-p\delta_{jk}$ (p now refers to the acoustic pressure field). Then the split equations (8) and (9) in Cartesian coordinates reduce to

$$\rho(a_q\partial_t+\omega_q)v_q=-\partial_q p,\quad q=1,2,3, \tag{95}$$

$$(a_q\partial_t+\omega_q)p^{(q)}=-\rho c^2\partial_q v_q,\quad q=1,2,3, \tag{96}$$

where $c=\sqrt{\lambda/\rho}$ is the acoustic wave velocity, and the repeated index q does not imply summation. These equations are the same as those given in Ref. 16 for a lossless medium. As noted in Ref. 16, there is no need to split the velocity field in Eq. (95).

A. 3-D cylindrical coordinates

From Eqs. (14)–(19) and (27), (28), the three-dimensional PML equations for scalar acoustic waves in cylindrical coordinates (r,θ,z) are

$$\rho(a_r\partial_t+\omega_r)v_r=-\partial_r p, \tag{97}$$

$$\rho(A_r\partial_t+\Omega_r)v_\theta=-\partial_\theta p, \tag{98}$$

$$\rho(a_z\partial_t+\omega_z)v_z=-\partial_z p, \tag{99}$$

$$(a_r\partial_t+\omega_r)p^{(r)}=-\rho c^2\partial_r v_r, \tag{100}$$

$$(A_r\partial_t+\Omega_r)p^{(\theta)}=-\rho c^2(v_r+\partial_\theta v_\theta), \tag{101}$$

$$(a_z\partial_t+\omega_z)p^{(z)}=-\rho c^2\partial_z v_z. \tag{102}$$

A total of six unknowns and six equations are needed, as for Cartesian coordinates.

B. 2-D polar coordinates

For 2-D polar coordinates (r,θ), the PML equations for scalar acoustic waves can be reduced from Eqs. (30)–(41) as

$$\rho(a_r\partial_t+\omega_r)v_r=-\partial_r p, \tag{103}$$

$$\rho(A_r\partial_t+\Omega_r)v_\theta=-\partial_\theta p, \tag{104}$$

$$(a_r\partial_t+\omega_r)p^{(r)}=-\rho c^2\partial_r v_r, \tag{105}$$

$$(A_r\partial_t+\Omega_r)p^{(\theta)}=-\rho c^2(v_r+\partial_\theta v_\theta). \tag{106}$$

That is, only four unknowns and four equations are required.

C. 2-D axisymmetric coordinates

For 2-D axisymmetric coordinates (r,z), the PML equations for scalar acoustic waves can be reduced from Eqs. (43)–(59) as

$$\rho(a_r\partial_t+\omega_r)v_r=-\partial_r p, \tag{107}$$

$$\rho(a_z\partial_t+\omega_z)v_z=-\partial_z p, \tag{108}$$

$$(a_r\partial_t+\omega_r)p^{(r)}=-\rho c^2\partial_r v_r, \tag{109}$$

$$(A_r\partial_t+\Omega_r)p^{(\theta)}=-\rho c^2 v_r, \tag{110}$$

$$(a_z\partial_t+\omega_z)p^{(z)}=-\rho c^2\partial_z v_z. \tag{111}$$

That is, five unknowns and five equations are required. It is interesting to note that as for the elastic waves, it takes more unknowns for the axisymmetric coordinates than for the polar coordinates.

D. 3-D spherical coordinates

Three-dimensional PML equations for scalar acoustic waves in spherical coordinates can be obtained from Eqs. (70)–(86) as

$$\rho(a_r\partial_t+\omega_r)v_r=-\partial_r p, \tag{112}$$

$$\rho(A_r\partial_t+\Omega_r)v_\theta=-\partial_\theta p, \tag{113}$$

$$\rho(A_r\partial_t+\Omega_r)v_\phi=-\frac{1}{\sin\theta}\partial_\phi p, \tag{114}$$

$$(a_r\partial_t+\omega_r)p^{(r)}=-\rho c^2\partial_r v_r, \tag{115}$$

$$(A_r\partial_t+\Omega_r)p^{(x)}=-\rho c^2\left[2v_r+\frac{1}{\sin\theta}\partial_\theta(v_\theta\sin\theta)+\frac{1}{\sin\theta}\partial_\phi v_\phi\right]. \tag{116}$$

It is not surprising to note that only five unknowns and five equations are required for spherical coordinates, compared to six in Cartesian coordinates.

V. NUMERICAL RESULTS

Although it is not the main purpose of this paper to discuss numerical results, two simple solutions are shown below to demonstrate the effectiveness of the PML absorber in cylindrical and spherical coordinates. For cylindrical coordinates, as the z direction is identical to that in Cartesian coordinates,[13] it is sufficient to show the results for polar coordinates without z variations. Numerical results will be shown for a medium with $c=\sqrt{(\lambda+2\mu)/\rho}=3000$ m/s, $s=\sqrt{\mu/\rho}=1500$ m/s, $\rho=2500$ kg/m^3, and an excitation frequency of 1 kHz. The interface between the regular medium and a PML medium is located at $r=a=5\lambda_c$ (where $\lambda_c=c/f$, $f=1$ kHz). Unless otherwise stated, the PML parameters inside the PML region are chosen as $\omega_r=2\pi f=2\pi\times10^3$ Hz. The choice of ω_r and ω_z is usually made so that they are several times the central frequency inside the frequency band of interest. This ensures that waves decay rapidly within few wavelengths inside the PML.

In polar (r,θ) coordinates, wave fields due to an arbitrary source can be expressed in terms of the superposition of cylindrical harmonics corresponding to the compressional (P) waves and shear (S) waves. Therefore, we study the behaviors of outgoing P and S waves near a radial interface at $r=a$ between a regular medium with Láme constants λ, μ and density ρ and an outer PML medium with the same physical properties and a complex e_r. In a regular homogeneous unbounded medium, without loss of generality the outgoing waves can be represented as Hankel functions $H_n^{(1)}(k_c r)$ and $H_n^{(1)}(k_s r)$, where $k_c=\omega\sqrt{\rho/(\lambda+2\mu)}$ and $k_s=\omega\sqrt{\rho/\mu}$ are the P and S wave numbers. In a PML medium with the same physical properties but with a complex func-

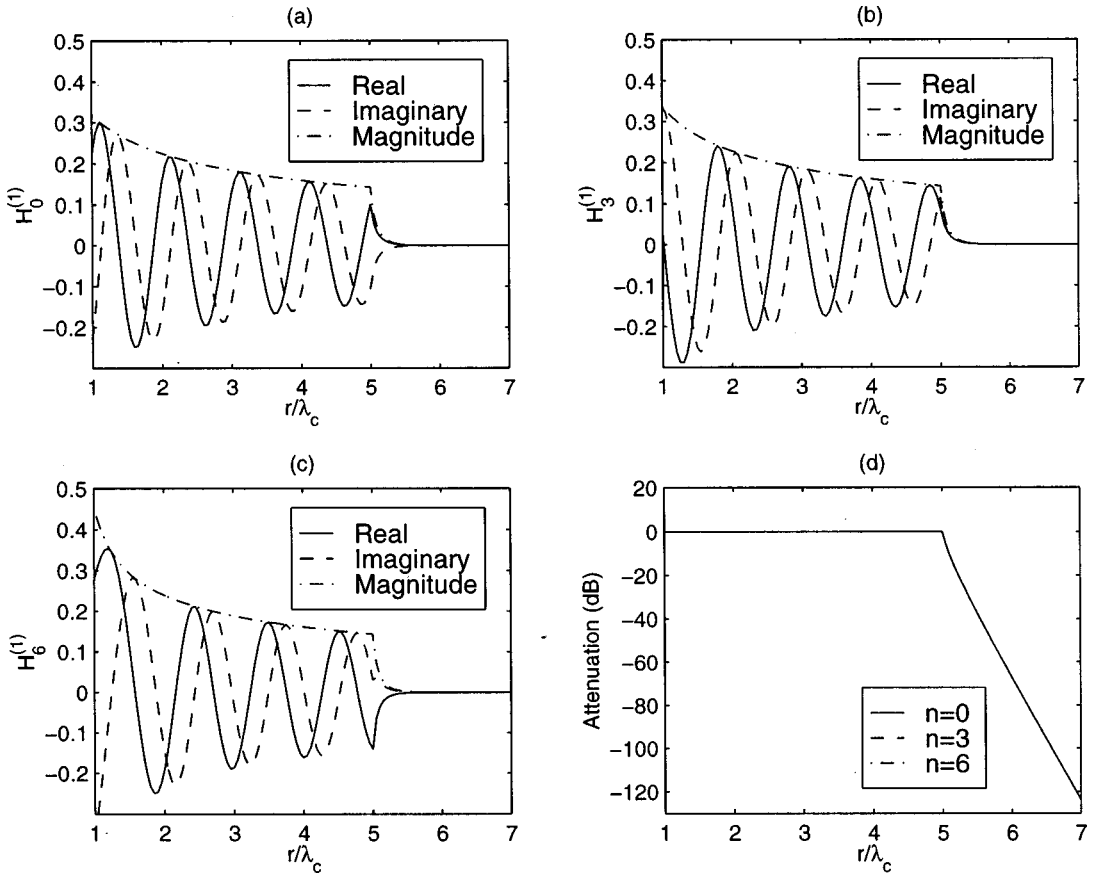


FIG. 1. Attenuation of compressional waves in an elastic PML medium in polar coordinates ($a_r=100$, $\omega_r=2\pi\times10^3$ Hz). (a) $H_0^{(1)}(k_c\tilde{r})$. (b) $H_3^{(1)}(k_c\tilde{r})$. (c) $H_6^{(1)}(k_c\tilde{r})$. (d) The attenuation ratio $H_n^{(1)}(k_c\tilde{r})/H_n^{(1)}(k_cr)$ for $n=0$, 3, and 6.

tion $e_r(r)$, these P and S waves become $H_n^{(1)}(k_c\tilde{r})$ and $H_n^{(1)}(k_s\tilde{r})$, respectively, where $\tilde{r}$ is given by Eq. (20). If we choose the complex $e_r(r)$ function as

$$e_r=\begin{cases}1, & \text{for } r<a\\ a_r+\dfrac{i\omega_r}{\omega}, & \text{for } r\geqslant a\end{cases},$$

where a_r and ω_r are positive constant, then inside the PML interface ($r=a$) is a regular medium, but outside is a PML medium. The effects of the outer PML medium on the absorption and reflection of outgoing waves can be studied by comparing $H_n^{(1)}(k_m\tilde{r})$ and $H_n^{(1)}(k_mr)$ (where $m=c,s$).

Figure 1(a)–(c) show the attenuation of P waves $H_n^{(1)}(k_c\tilde{r})$ inside the PML region for $n=0$, 3, and 6. As expected, the P waves attenuate rapidly inside the PML region ($r\geqslant a$). Figure 1(d) display the ratio of the P wave $H_n^{(1)}(k_c\tilde{r})$ in this two-layer medium and $H_n^{(1)}(k_cr)$ for a regular unbounded medium. The fact that this ratio is 1 for $r<a$ in Fig. 1(d) confirms that there are no reflections at the PML interface. Furthermore, the attenuation ratios for all these n values are the same. The effect of a_r on the attenuation is shown in Fig. 2. It is seen that a large a_r enhances the attenuation. Similar observations can be made for the S waves in this two-layer medium, and for waves in spherical coordinates.

From the above example for propagating P waves, it is noted that although choosing $a_r>1$ can enhance the attenuation of waves in the PML region, the effect is not very significant. However, for evanescent waves, a_r has a more drastic effect. This can be seen from Fig. 3(a)–(d) for evanescent P waves in spherical coordinates with $a_r=10$. Comparison of $a_r=10$ and $a_r=1$ in Fig. 4 confirms that increasing a_r can significantly increase the attenuation rate in the PML region for evanescent waves. On the other hand, the effect of ω_r on the absorption is more obvious: A larger ω_r

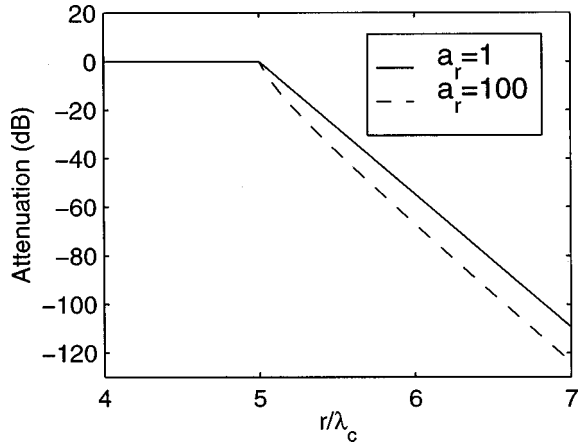


FIG. 2. Effect of a_r on the absorption of compressional waves for $n=0$ in polar coordinates.

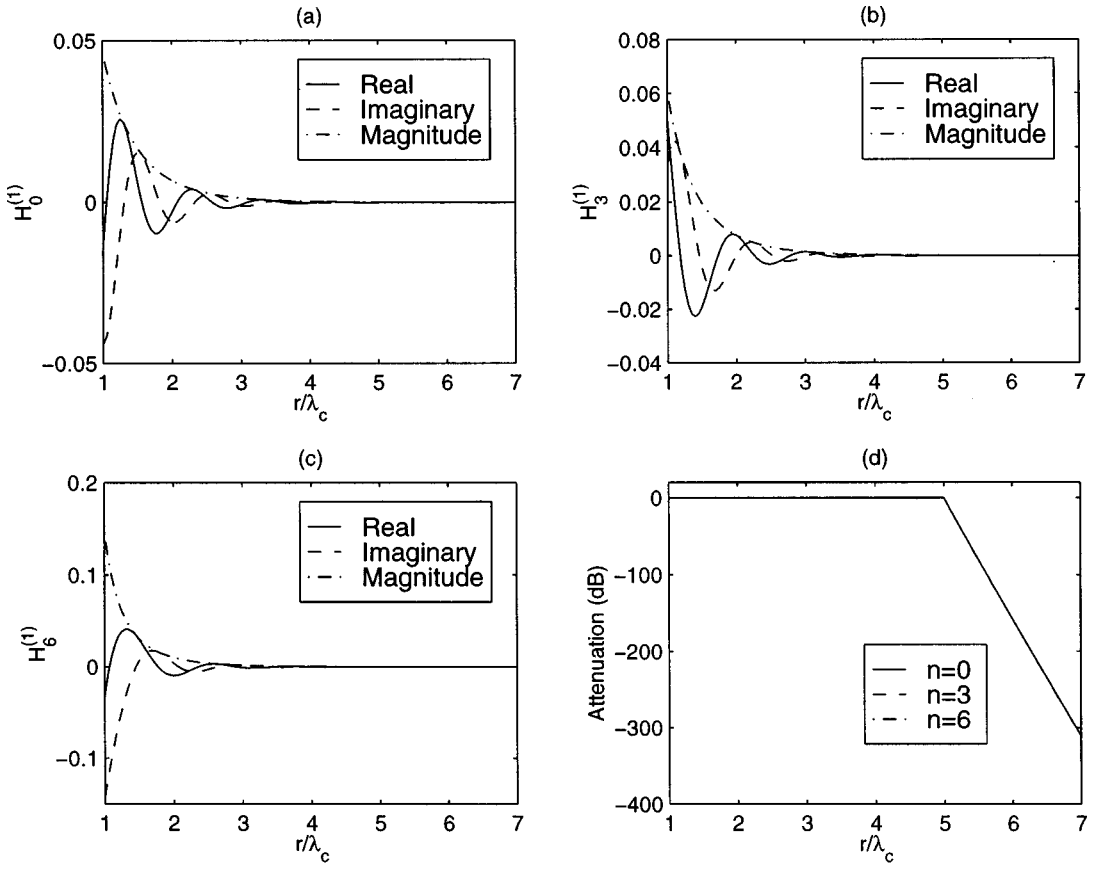


FIG. 3. Absorption of attenuative (evanescent) compressional waves by an elastic PML medium in spherical coordinates ($a_r=10$, $\omega_r=2\pi\times10^3$ Hz), where $\tilde{k}_c=(1+0.2i)k_c/\sqrt{1.04}$. (a) $h_0^{(1)}(\tilde{k}_c\tilde{r})$. (b) $h_3^{(1)}(\tilde{k}_c\tilde{r})$. (c) $h_6^{(1)}(\tilde{k}_c\tilde{r})$. (d) The attenuation ratio $h_n^{(1)}(\tilde{k}_c\tilde{r})/h_n^{(1)}(\tilde{k}_c r)$ for $n=0$, 3, and 6.

will introduce more rapid attenuation to the outgoing waves.

In the above examples, for convenience we have chosen e_r to be a step function. In numerical solutions using finite-difference time-domain (FDTD) or pseudospectral time-domain (PSTD) methods, a tapered PML profile should be chosen to reduce the reflections due to discretization. Furthermore, a truncating boundary has to be placed in the PML region. These issues are similar to those for Cartesian coordinates.[13,16]

We show an example of the PML applications in a finite-difference time-domain method for the axisymmetric case encountered in borehole acoustic measurements. The source time function is a Blackman–Harris window function with a central frequency of $f_c=5$ kHz. The PML variables are chosen such that $a_r=a_z=1$, and ω_r and ω_z have a linear profile of 10 cells at the outer boundaries with $\omega_{r,\max}=\omega_{z,\max}=16\pi f_c$.

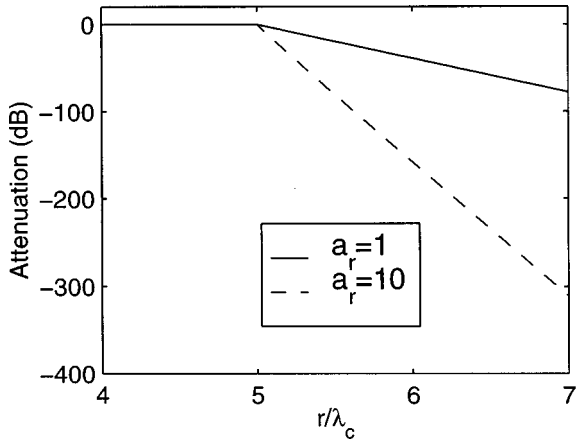


FIG. 4. Effect of a_r on the PML absorption of evanescent compressional waves for $n=0$ in spherical coordinates.

The first FDTD example is for a homogeneous solid with $c=3000$ m/s, $s=1500$ m/s, $\rho=2000$ kg/m^3. A grid of $N_r\times N_z=61\times241$ with $\Delta r=\Delta z=1.2$ cm, $\Delta t=1.5\,\mu$s is used to discretize the problem. A monopole ring source is located at cell index $(j,k)=(2,30)$. A reference case with a much larger grid is compared to characterize the PML reflection errors. Within the time window of interest, there are no reflections from the outer boundary in the reference case. Figure 5(a) shows the distribution of $\tau_{\theta\theta}$ along the radial locations $(j,15)$ where $j=0,...,60$. Excellent agreement between the two sets of results is observed in the non-PML region of interest. Within the radial PML region, the field decays rapidly. The small reflection errors are about -50 dB as shown in the inset. Figure 5(b) shows the waveform at (47, 15) as a function of time. Again, compared with the reference case, the reflection error is less than 0.5%. Further reduction of the reflection errors is possible.

We show one application of the FDTD method in the simulation of the borehole acoustic measurements. A borehole (radius 10 cm) with a fluid of $c=1500$ m/s, $\rho=1000$ kg/m^3 is surrounded by a solid of $c=3000$ m/s, $s=1100$ m/s, $\rho=2200$ kg/m^3. The source is the same as in the

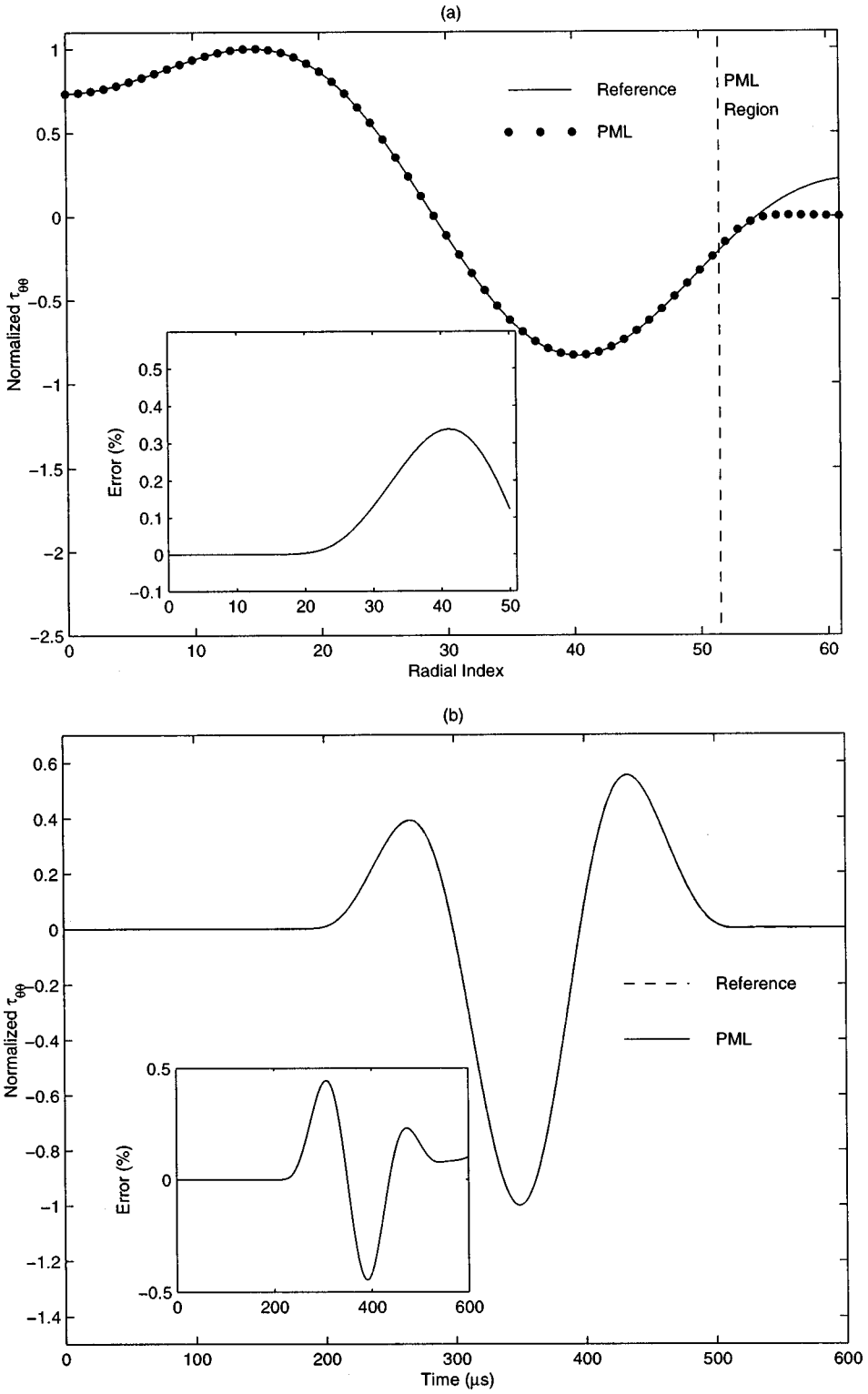


FIG. 5. PML reflection errors for the FDTD result. (a) The distribution of $\tau_{\theta\theta}$ along the radial direction ($z = 15$ cm from the source). (b) The waveform at $z = 15$ cm, $r = 48$ cm (3 cells from the PML interface). The insets display the relative errors. The reference case has a large grid and does not have reflections within the time window.

last example and is located at (2, 20), while an array of receivers are located at $(2,5k+25)$ where $k=1,...,17$. The problem is discretized with a grid of $N_r \times N_z = 61 \times 121$ with $\Delta r = \Delta z = 1$ cm, $\Delta t = 1.25\,\mu$s. In Fig. 6 we compare the FDTD results with the real-axis integration (RAI) method.[25] Again, excellent agreement is observed.

The PML method presented here is also useful for full three-dimensional FDTD and pseudospectral time-domain (PSTD) methods.[24] These applications will be investigated in our future research.

VI. CONCLUSIONS

A perfectly matched layer (PML) is developed for elastic waves in cylindrical and spherical coordinates using the improved scheme of complex coordinates. With the introduction of the integrated complex stretching variables, the cylindrical and spherical PML schemes require no more unknown split field components than in Cartesian coordinates. The formulas for the special cases with 2-D polar and axisymmetric coordinates and for the scalar acoustic waves are

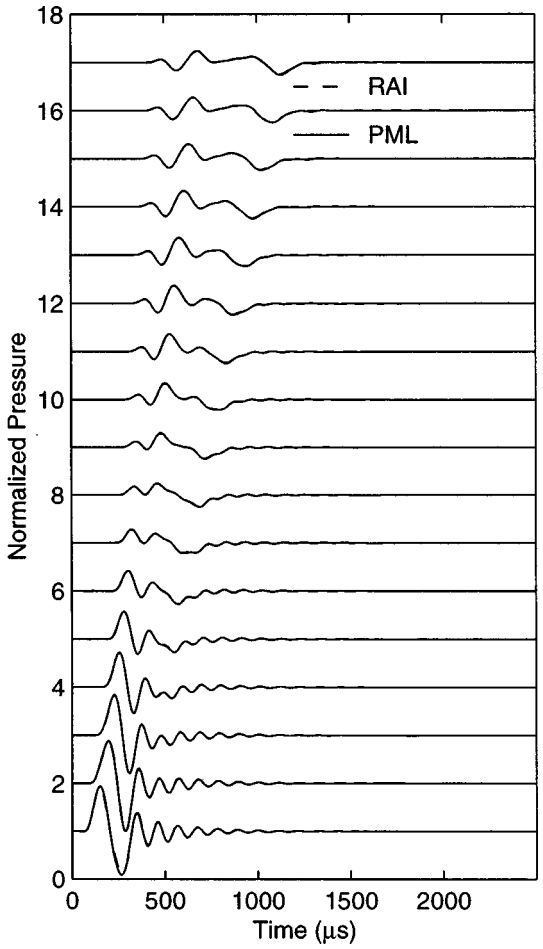


FIG. 6. Comparison of the FDTD results with the real-axis integration method for a borehole acoustic measurement.

also given. The PML absorbing boundary condition is important for the truncation of unbounded domains in numerical methods for partial differential equations. Numerical results confirm the efficacy of the PML absorbing boundary condition.

ACKNOWLEDGMENTS

The author would like to thank Dr. Raymond Nagem and an anonymous reviewer for their useful comments to improve the presentation of the manuscript. This work was supported by U.S. EPA through a PECASE Grant No. CR-825-225-010, and by the NSF through a CAREER Grant No. ECS-9702195.

[1] C. Cerjan, D. Kosloff, R. Kosloff, and M. Reshef, ''A nonreflecting boundary condition for discrete acoustic and elastic wave equations,'' Geophysics **50**, 705–708 (1985).
[2] A. R. Levander, ''Use of the telegraphy equation to improve absorbing boundary efficiency for fourth-order acoustic wave finite difference schemes,'' Bull. Seismol. Soc. Am. **75**, 1847–1852 (1985).
[3] R. Clayton and B. Engquist, ''Absorbing boundary conditions for acoustic and elastic wave equations,'' Bull. Seismol. Soc. Am. **67**, 1529–1540 (1977).
[4] E. L. Lindman, ''Free-space boundary conditions for the time dependent wave equation,'' J. Comput. Phys. **18**, 66–78 (1975).
[5] C. J. Randall, ''Absorbing boundary condition for the elastic wave equation: Velocity-stress formulation,'' Geophysics **54**, 1141–1152 (1989).
[6] Z. P. Liao, H. L. Wong, B. P. Yang, and Y. F. Yuan, ''A transmitting boundary for transient wave analysis,'' Sci. Sinica A **27**, 1063–1076 (1984).
[7] R. L. Higdon, ''Numerical absorbing boundary conditions for the wave equation,'' Math. Comput. **49**, 65–90 (1987).
[8] C. Peng and M. N. Toksöz, ''An optimal absorbing boundary condition for finite difference modeling of acoustic and elastic wave propagation,'' J. Acoust. Soc. Am. **95**, 733–745 (1994).
[9] J. B. Schneider and O. M. Ramahi, ''The complementary operators method applied to acoustic finite-difference time-domain simulations,'' J. Acoust. Soc. Am. **104**, 686–693 (1998).
[10] Y. H. Chen, W. C. Chew, and Q. H. Liu, ''A three-dimensional finite difference code for the modeling of sonic logging tools,'' J. Acoust. Soc. Am. **103**, 702–712 (1998).
[11] J.-P. Berenger, ''A perfectly matched layer for the absorption of electromagnetic waves,'' J. Comput. Phys. **114**, 185–200 (1994).
[12] W. C. Chew and W. H. Weedon, ''A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates,'' Microwave Opt. Technol. Lett. **7**, 599–604 (1994).
[13] W. C. Chew and Q. H. Liu, ''Perfectly matched layers for elastodynamics: A new absorbing boundary condition,'' J. Comput. Acoust. **4**, 72–79 (1996).
[14] F. D. Hastings, J. B. Schneider, and S. L. Broschat, ''Application of the perfectly matched layer (PML) absorbing boundary condition to elastic wave propagation,'' J. Acoust. Soc. Am. **100**, 3061–3069 (1996).
[15] J. G. Maloney and K. E. Cummings, ''Adaptation of FDTD techniques to acoustic modeling,'' in 11th Annual Review of Progress in Applied Computational Electromagnetics, Vol. 2 (Monterey, CA), pp. 724–731 (1995).
[16] Q. H. Liu and J. Tao, ''The perfectly matched layer (PML) for acoustic waves in absorptive media,'' J. Acoust. Soc. Am. **102**, 2072–2082 (1997).
[17] X. Yuan, D. Borup, J. W. Wiskin, M. Berggren, R. Eidens, and S. A. Johnson, ''Formulation and validation of Berenger's PML absorbing boundary for the FDTD simulation of acoustic scattering,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **44**, 816–822 (1997).
[18] Q. H. Liu, F. Daube, C. Randall, E. Schoen, H. Liu, and P. Lee, ''A three-dimensional finite difference simulation of sonic logging,'' J. Acoust. Soc. Am. **100**, 72–79 (1996).
[19] Q. H. Liu and J. Q. He, ''Quasi-PML for waves in cylindrical coordinates,'' New Mexico State University Technical Report, July 1997. Also in Microwave Opt. Technol. Lett., **19**, 107–111, 1998.
[20] W. C. Chew, J. M. Jin, and E. Michelssen, ''Complex coordinate system as a generalized absorbing boundary condition,'' Microwave Opt. Technol. Lett. **15**, 363–369 (1997).
[21] F. L. Teixeira and W. C. Chew, ''PML-FDTD in cylindrical and spherical grids,'' IEEE Microwave Guid. Wave Lett. **7**, 285–287 (1997).
[22] B. Yang, D. Gottlieb, and J. S. Hesthaven, ''Spectral simulations of electromagnetic wave scattering,'' J. Comput. Phys. **134**, 216–230 (1997).
[23] J. He and Q. H. Liu, ''A nonuniform cylindrical FDTD algorithm with improved PML and quasi-PML absorbing boundary conditions,'' IEEE Trans. Geosci. Remote Sens. to appear in March (1998).
[24] Q. H. Liu, ''The PSTD algorithm for acoustic waves in inhomogeneous, absorptive media,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **45**, 1044–1055 (1998).
[25] C. C. Lu and Q. H. Liu, ''A three-dimensional dyadic Green's function for elastic waves in multilayer cylindrical structures,'' J. Acoust. Soc. Am. **98**, 2825–2835 (1995).

Temperature dependence of collective phonon relaxation time and acoustic attenuation in pure GaAs

S. D. Lambade
Department of Physics, Y. C. College of Engineering, Wanadongri, Nagpur 441 110, India

G. G. Sahasrabudhe
Department of Physics, R. K. N. Engineering College, Nagpur 440 013, India

(Received 15 July 1997; revised 12 November 1998; accepted 14 November 1998)

In this paper, the temperature dependence of collective phonon-relaxation time in pure GaAs in the range 80–300 K was investigated using a refined numerical approach in which the interaction of the acoustic wave with the complete spectrum of phonon modes is taken into consideration. These are further used to evaluate the acoustic attenuation for longitudinal waves along the [110] and [111] directions and for shear waves along the [110] direction with polarizations along [110] and [001]. The calculations are performed using Mason's pure-mode scheme, Mason's theory with refinement suggested by Merkulov *et al.*, and also using Nava's modified formulation of Woodruff's theory. The results obtained using this computational scheme show better agreement with the experiment than those obtained using Mason's pure-mode scheme in most of the cases studied. © *1999 Acoustical Society of America.* [S0001-4966(99)01503-9]

PACS numbers: 43.25.Ed [MAB]

INTRODUCTION

An acoustic wave interacts with thermal lattice waves by virtue of the anharmonic nature of the interatomic forces in a solid. This interaction is dissipative. Absorption of a sound wave due to its interaction with thermal phonons has been reviewed by Maris.[1] This study is generally divided into two theoretical regimes: the Landau–Rumer, and Akhiezer regimes.

The acoustic wave interacts directly with individual lattice phonons if $\Omega\tau \gg 1$, where Ω is the angular frequency of the sound wave and τ the relaxation time of the thermal phonons. This corresponds to the case when the wavelength λ of the wave is much less than the mean free path l of the thermal phonons. This condition is satisfied for ultrasonic waves only at the lowest temperatures and highest frequencies. The condition $T \ll \theta_{D'}$, where θ_D is the Debye temperature, can be safely assumed in this regime. Landau and Rumer[2] have given a quantum mechanical treatment of the acoustic losses in this regime. The sound waves and thermal waves are both treated microscopically, i.e., as phonons, by Landau and Rumer. In the range $\Omega\tau \ll 1$ (which is equivalent to the condition $\lambda \gg 1$), a condition realized at higher temperatures, the uncertainty in the energy of the thermal phonons is greater than the energy of the phonons being propagated, and the effect of the sound wave on the entire assembly of thermal phonons must be considered. In this regime, the first theory was proposed by Akhiezer,[3] who treated the sound wave macroscopically. A more complete development of this idea was given by Woodruff and Ehrenreich,[4] who utilize the complete Boltzmann equation to calculate the steady-state distribution of thermal phonons and the sound attenuation.

The study of temperature dependence of acoustic attenuation in semiconductors in the Akhiezer regime has stimulated much interest among researchers in the last three decades.[5–10] The pioneering contribution to this study was made by Mason and Bateman, who measured the temperature variation of attenuation in Si and Ge.[11,12] They also suggested a method of calculation involving the nonlinearity constant, D, which is a function of second- and third-order elastic constants (SOEC and TOEC). Due to the lack of experimental data of temperature dependence of TOEC at that time, D was considered independent of temperature by Mason and Bateman. When the experimental data on temperature variation of TOEC was reported by Breazeale *et al.*,[13] temperature dependence of D was studied by Rajagopalan *et al.*[14–16] and Joharapurkar,[17] who found that the incorporation of the temperature dependence of D in the calculation of attenuation resulted in improvement in agreement with the experimental data.

Nava *et al.*[8] applied the more rigorous Woodruff and Ehrenreich formulation,[4] modified slightly to include the dispersion and anisotropy of the various thermal phonon parameters, to extract the temperature dependence of the effective ultrasonic Grüneisen parameter (UGP) in Si and Ge using the experimental temperature-dependent attenuation data. Recently, Ilisavskii *et al.*[18] have reported the experimental study of attenuation in Si, covering both the regimes of acoustic damping from which they could determine the temperature dependence of the collective phonon-relaxation time.

Among the III–V semiconducting compounds, GaAs has been the most widely studied for its attenuating characteristics.[19–25] However, the theoretical investigations of the temperature dependence of attenuation in GaAs have not been reported so far due to lack of experimental data on temperature dependence of TOEC. Recently, Joharapurkar *et al.*[26] have reported experimental determination of the temperature dependence of the complete set of TOEC for GaAs in the range 77–300 K. We use these data to present a comprehensive calculation of acoustic damping in GaAs.

 0001-4966/99/105(4)/2085/8/$15.00 © 1999 Acoustical Society of America

In this paper, we suggest a new computational approach to determination of collective phonon-relaxation time. The approach is based on the work of Ilisavskii *et al.*[18] and Hrivnak *et al.*[21] These calculations are performed in the temperature range 80–300 K. These relaxation times are further used to estimate the temperature dependence of attenuation for longitudinal waves along the [110] and [111] directions and for shear waves along the [110] direction (polarized along $[1\bar{1}0]$ and [001]).

The calculations are performed using the computational scheme developed by us,[27] and also using Mason's scheme restricted only to pure modes. All previous workers (in the references cited above) have assumed Mason's scheme in performing these calculations, though serious objections have been raised by Barrett and Holland[6] against Mason's theory, and many workers[28–32] including Mason himself[1,11,12] have stressed the need for refinement in his format of calculation.

The computer program developed by us considers the interaction of the acoustic wave with the entire spectrum of phonon modes, and facilitates the more complete calculation of Mason's theory using Merkulov *et al.*'s correction factor[29] and Nava's modified formulation of Woodruff's theory. Earlier, the program had been used with considerable success in accounting for the temperature dependence of attenuation in Si, Ge,[27,33,34] NaCl, NaF, and LiF.[35–37]

I. THEORY

A. Relaxation time

In the calculation of attenuation, the relaxation time for transverse acoustic waves is usually assumed to be close to the thermal relaxation time obtained using the expressions

$$\tau_{\text{th}}=3\kappa/C_V c^2, \quad c^2=(2V_s^{-3}+V_l^{-3})/3, \tag{1}$$

where V_l and V_s are the longitudinal and shear-wave velocities, respectively.

A few years back, Ilisavskii *et al.*[18] reported the temperature dependence of the experimental collective phonon-relaxation time for Si. These data are obtained from identification of kinks in the frequency-attenuation plots at various temperatures with the transition from attenuation governed by the Akhiezer mechanism ($\Omega\tau \ll 1$) to that governed by the Landau–Rumer mechanism ($\Omega\tau > 1$). Thus, at these kinks, the condition $\Omega\tau=1$ is satisfied. The value of τ found in this way is an averaged one for the whole system of thermal phonons. The experimental relaxation times reported by Ilisavskii *et al.* agree with Eq. (1) in the low-temperature range. However, in the high-temperature range, relaxation times calculated using Eq. (1) show a much faster decrease with rise in temperature compared to the experimental τ. Ilisavskii *et al.* attributed this to the fact that the main contribution to the thermal conductivity in Si comes from transverse thermal phonons, while C_v in Eq. (1) is generally taken to be the *total* specific heat. The fact that the bulk of the thermal conductivity is contributed mainly by transverse phonons has been observed for GaAs, among other compounds.[38–41]

In our computational scheme, it is possible to calculate the contributions from various phonon branches to the total specific heat. Thus, the correct calculation of τ is possible with our program. The discrepancy between the experimental and the calculated relaxation times can be removed by expressing the collective phonon-relaxation time in a more appropriate form as

$$\tau=3\kappa/C_T c^2, \tag{2}$$

where C_T is part of the specific heat that corresponds only to the transverse phonon branches.

The total thermal conductivity of a solid can be expressed as

$$\kappa=\kappa_L+\kappa_{T_1}+\kappa_{T_2}. \tag{3}$$

Hrivnak and Kovar[21] have used the substitution

$$\frac{\kappa}{c^2}=\frac{\kappa_L}{\overline{V}_L^2}+\frac{\kappa_{T_1}}{\overline{V}_{T_1}^2}+\frac{\kappa_{T_2}}{\overline{V}_{T_2}^2}, \tag{4}$$

in the expression for attenuation coefficient used by them. In these equations, κ_L, κ_{T_1}, and κ_{T_2} are contributions to the lattice thermal conductivity from longitudinal, slow, and fast shear phonons, respectively, whose Debye average velocities are $\overline{V}_L$, $\overline{V}_{T_1}$, and $\overline{V}_{T_2}$. For our computational purpose, we consider this substitution as the definition of c, instead of regarding c as Debye average velocity in the usual sense. Thus, we can write

$$\frac{1}{c^2}=\frac{1}{\overline{V}_{T_1}^2}\left[\frac{1+(\kappa_{T_2}\overline{V}_{T_1}^2/\kappa_{T_1}\overline{V}_{T_2}^2)}{1+(\kappa_{T_2}/\kappa_{T_1})}\right], \tag{5}$$

where the approximation[38–41] $\kappa \approx \kappa_{T_1}+\kappa_{T_2}$ is used. Assuming further that the relaxation times for slow and fast transverse modes are the same, and that the kinetic equation $\kappa=1/3\langle CV^2\rangle\tau$ applies separately to thermal conductivity contributions by slow and fast transverse modes, we can write

$$\frac{1}{c^2}=\frac{1}{\overline{V}_{T_1}^2}\left[\frac{1+\left[\frac{\langle C_{T_2}V_{T_2}^2\rangle}{\langle C_{T_1}V_{T_1}^2\rangle}\right]\left[\frac{\overline{V}_{T_1}^2}{\overline{V}_{T_2}^2}\right]}{1+\left(\frac{\langle C_{T_2}V_{T_2}^2\rangle}{\langle C_{T_1}V_{T_1}^2\rangle}\right)}\right]. \tag{6}$$

The computation of the various averages in Eq. (6) is possible using our program over the complete spectrum of phonon modes. Thus, the evaluation of the collective phonon-relaxation time is facilitated in our program using Eqs. (2) and (6).

B. Acoustic attenuation

The absorption of acoustic waves in the range $\Omega\tau \ll 1$ is mainly a result of the interaction of acoustic phonons with thermal phonons. The dominant source of attenuation in insulators is the Akhiezer loss, which is due to the inequilibrium in the thermal phonon distribution arising from the passage of the sound wave. The perturbed phonon distribution

relaxes towards equilibrium via anharmonic phonon–phonon collisions in an entropy-producing process, which damps the wave.

Akhiezer's original theory was modified first by Woodruff and Ehrenreich and then by Mason and Bateman, using fundamentally different approaches. Woodruff's expression for attenuation is a function of mode-specific heats, and involves complex averages of the anharmonic parameters, which describe the strain dependence of the lattice vibrational frequencies over the Brillouin zone. Mason and Bateman suggested another derivation in the anisotropic continuum limit, their expression for attenuation coefficient is

$$A(\mathrm{Np/cm}) = \left[\frac{E_0 D}{6\rho V^3}\right]\left[\frac{\Omega^2\tau}{1+\Omega^2\tau^2}\right], \tag{7}$$

where E_0 is the thermal energy per unit volume, ρ is the density, V is the mode velocity, and the nonlinearity constant, D, is given by

$$D = 3\left[3\sum \frac{[\gamma(i)]^2}{n} - \langle\gamma\rangle^2 C_v T/E_0\right], \tag{8}$$

where T is the absolute temperature, and the thermal Grüneisen constant is expressed as the average $\langle\gamma\rangle$ of the mode Grüneisen numbers $\gamma(i)$. Mason and Bateman prescribe the use of Brugger's expressions[42] for evaluating the mode Grüneisen numbers. In Brugger's notation, these numbers are given by

$$\gamma_i^{jk} = (1/2W)[2WU_jU_k(C_{jkmn} + C_{jkmunu}U_uU_v)N_mN_n]$$

and

$$\gamma_i^{j'k'} = \left[\frac{\partial x_{j'}}{\partial x_j}\right]\left[\frac{\partial x_{k'}}{\partial x_k}\right]\gamma_i^{jk}, \tag{9}$$

with

$$W = C_{munv}N_mN_nU_uU_v, \tag{10}$$

where Ns and Us are the direction cosines of the propagation direction and the polarization direction, respectively, in the acoustic mode i, Cs are the elastic constants, $\gamma_i^{j'k'}$ are the Grüneisen numbers along an axis other than cube axes, and $(\partial x_{j'}/\partial x_j)$ are the direction cosines between the new axes and the old axes.

Earlier, we showed[27] that the refinement suggested by Merkulov *et al.*[29] in Mason's theory can be incorporated in our program through the use of the correction factor

$$M_i = 1 - 1/\{3I'(\omega_{gi})[\exp(\hbar\omega_{gi}/K_\mathrm{B}T) - 1]\}, \tag{11}$$

where K_B is the Boltzmann's constant, ω_{gi} is the cutoff frequency supported by the lattice, and the integral $I'(\omega_{gi})$ is given by

$$I'(\omega_{gi}) = \int_0^1 d\mu\,\mu^3/[\exp(\hbar\omega_{gi}\mu/K_\mathrm{B}T) - 1]. \tag{12}$$

The modified formulation of Woodruff's theory recently proposed by Nava *et al.*[8] is based on the use of the effective ultrasonic Grüneisen parameter (UGP). Their expression for the attenuation coefficient for a wave of polarization j and propagation direction **K** is,

$$A_j(\mathrm{Np/cm}) = \left[\frac{3\Omega^2\kappa_\mathbf{K}T}{\rho V_j^3 c^2}\right]\Gamma^2, \tag{13}$$

where $\kappa_\mathbf{K}$ is the thermal conductivity along **K**, V_j is the appropriate sound-wave velocity, and Γ^2 is the effective UGP. This parameter is an average of GP for each phonon mode $(\mathbf{q},i)$ weighted by their thermal conductivity $\kappa_\mathbf{K}(\mathbf{q},i)$ and is formally given by the expression

$$\Gamma_j^2 = \frac{\sum_{\mathbf{q},i} \Gamma_{ij}^2(\mathbf{q},i)\kappa_\mathbf{K}(\mathbf{q},i)}{\sum_{\mathbf{q},i} \kappa_\mathbf{K}(\mathbf{q},i)}, \tag{14}$$

with

$$\Gamma_{ij}^2 = (c^2/\beta_i^2)\left[\gamma_{ij}^2 - \langle\langle\gamma_{ij}\rangle\rangle\gamma_{ij}\left[1 - \frac{\beta_i^2}{V_j^2}\right]\delta_{jL}\right], \tag{15}$$

and

$$\kappa_\mathbf{K}(\mathbf{q},i) = \tfrac{1}{3}C_i\beta_i^2\tau_i. \tag{16}$$

Here, all the properties of the thermal phonon of a given branch i are assumed to depend on the wave vector **q**. β_i is the component of the phonon group velocity $\mathbf{V}_i$ along **K**, C_i is the mode-specific heat, γ_{ij} is the mode-generalized GP, $\langle\langle\gamma_{ij}\rangle\rangle$ is the specific heat-weighted average for all modes in the Brillouin zone, and δ_{jL} is the Kronecker's delta.

Nava *et al.* assume an anisotropic elastic continuum model in which the **q** dependence of the phonon properties is neglected and γ_{ij} are simply given by linear combinations of second- and third-order elastic constants (SOEC and TOEC). They further assume a constant phonon-relaxation time deduced from the thermal conductivity using Eq. (1). Under these approximations, Eq. (15) reduces to

$$\Gamma_j^2 = \langle\gamma_{ij}^2\rangle - \langle\gamma_{ij}\rangle[\langle\gamma_{ij}\rangle - \langle\gamma_{ij}\beta_i^2/V_j^2\rangle]\delta_{jL}, \tag{17}$$

where

$$\langle\alpha_{ij}\rangle = \sum_i C_i\alpha_{ij}/C. \tag{18}$$

Assuming further that all phonon modes contribute equally to the specific heat, Nava *et al.* approximate the averages indicated in Eq. (18) by simple numerical averages; for example

$$\langle\gamma_{ij}\rangle = \sum_i n_i\gamma_{ij}/N, \tag{19}$$

where n_i is the number of pure thermal modes coupled to the sound wave through γ_{ij}, and N is the total number of modes in the forward hemisphere. Mason's scheme of counting these modes is followed.

II. COMPUTATION

The computer program developed by us is based on the Brugger–Fritz[43] scheme of integration over the length of the wave vector followed by double angular integration over all directions. In this scheme, the continuum model is assumed. This means that the excitation of optic modes can be ne-

glected and the branch index for the mode takes only three values. The acoustic modes obey Debye distribution function and the maximum value of the length of the wave vector in any direction is the Debye radius $(6\pi^2/V_0)^{1/3}$, where V_0 is the volume of the primitive unit cell. The volume of the Debye sphere equals that of the first Brillouin zone. We have checked the consistency of results using different single- and multiple-integration routines.[44]

In our program, the eigenvalue equation $\rho V^2 U_i = C_{kilj} N_k N_l U_j$ is solved using the matrix diagonalization routine to get the three polarization eigenvectors for a particular direction of propagation. Except for pure-mode directions, which are a zero-measure subset among all directions, these modes are neither pure longitudinal nor pure transverse. However, the polarizations are still dominantly longitudinal or transverse. One can, therefore, estimate the fractional contribution of dominantly transverse modes to the specific heat.

The values of total specific heat obtained using our program showed some disagreement with those calculated using de Launay's technique.[45] Therefore, we used the product of de Launay specific heat and the ratio of partial specific heat due to transverse phonons to the total specific heat in evaluating C_T in Eq. (2).

The flexibility of our program allows us to calculate the averages of first and second power of mode Grüneisen numbers using one of the following procedures by selecting appropriate weights.

(a) In this scheme, the mode Grüneisen numbers are weighted by the number of thermal modes coupled to the sound wave through respective γs.

(b) The Merkulov correction factor expressed in Eq. (11) has been incorporated in our computational scheme. Thus, the mode γs are additionally weighted by this factor in computation of $\langle\gamma\rangle$ and $\langle\gamma^2\rangle$. The refinement suggested by Merkulov *et al.* could be achieved in our program through the numerical evaluation of the integral $I'(\omega_{gi})$ of Eq. (12). Thus, the explicit estimation of the effect of dependence of the integral

$$I(\omega_{gi}) = \int_0^{\omega_{gi}} \omega^3 \, d\omega / [\exp(\hbar\omega/K_B T) - 1]$$

on ω_{gi} is possible in our program. This removes the serious difficulty pointed out by Barrett and Holland in Mason's assumption that the integral $I(\omega_{gi})$ is independent of ω_{gi}, and thus leads to a more correct calculation of Mason's theory.

(c) The averages of mode γs indicated in Nava *et al.*'s formulation [Eqs. (17) and (18)] are defined by considering the mode-specific heats as the weights. However, the calculations reported by Nava *et al.* have been performed using approximate simple numerical averages, following Mason's scheme. Our program facilitates a more complete calculation of Nava *et al.*'s modified formulation of Woodruff's theory by performing the specific-heat weighted averages of mode γs over the entire spectrum of thermal phonon modes. Moreover, the calculation of the average $\langle \gamma_{ij} \beta_i^2 / V_j^2 \rangle$, is performed correctly in our program by evaluating the component β_i of the phonon group velocity along the sound-wave propagation direction for each phonon mode. β_i has not been used in this sense by Nava *et al.*, although it is defined like this.

In Mason's scheme, the numerical values of γ_i^j for pure modes are calculated using expressions tabulated by Mason and Bateman[5,11,12] and Merkulov *et al.*[29] The averages $\langle\gamma\rangle$ and $\langle\gamma^2\rangle$ are then obtained by taking the weighted averages of these γs over 39 (longitudinal) and 18 or 20 (shear) waves.

For calculation of total attenuation for longitudinal waves using Mason's theory, the contribution due to thermoelastic loss is evaluated separately using the expression

$$A(\text{Np/cm}) = \Omega^2 \langle\gamma\rangle \kappa T / 2\rho V_1^5. \tag{20}$$

Thermoelastic loss, which is usually very small for dielectrics (less than 4%), arises due to the flow of thermal energy from the compressed hotter part to the expanded cooler part associated with the compressional wave. There is no thermoelastic loss for transverse waves. In Nava *et al.*'s scheme, the contribution of thermoelastic loss is included in the calculation of UGP [Eq. (17)].

In our program, the number of directions for which γ_i^j are evaluated is determined by a convergence criterion for the numerical approximation of the appropriate integral. The minimum number of directions considered in the present study was 1600. For shear waves, $\langle\gamma\rangle$ was found to be consistently less than 10^{-6} with the imposed convergence criterion in all the cases studied. This points to the essential correctness of the numerical approximation to the full integral.

In our calculation of attenuation, we use the relaxation times obtained using our refined computational approach. The Debye average velocities and the fraction of specific heat contributed only by the transverse thermal phonons were evaluated at different temperatures using our computer program. The data of temperature dependence of thermal conductivity used in this calculation are those reported by Holland *et al.*[46]

We have taken into account the temperature dependence of each parameter in Eqs. (7), (13), and (20) in performing the calculation of temperature dependence of attenuation. Temperature dependence of density is calculated using the room-temperature density and the data on the temperature-dependent thermal expansion coefficient.[47] The experimental data of temperature-dependent SOEC reported by Cottam *et al.*[48] are used. The temperature-dependent TOEC data used are those reported by Joharapurkar *et al.*[26] The mode velocities along different directions are obtained using the temperature-dependent SOEC, density data, and the relevant expressions.[49] The Debye characteristic temperature, which is required in the estimation of C_v and E_0, is evaluated at different temperatures using de Launay's technique.[45]

III. RESULTS AND DISCUSSION

Experimental determination of temperature variation of the collective phonon-relaxation time has so far been reported only for Si.[18] Figure 1 presents a comparison of our computed relaxation times, obtained with the same data used earlier,[27,34] with the experimental data of temperature depen-

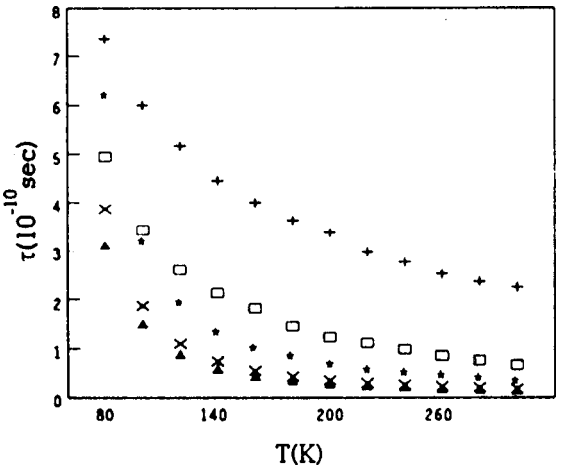


FIG. 1. Temperature dependence of relaxation time for Si. + and □: experimental values reported by Ilisavskii *et al.*[18] for longitudinal waves along [100] and [110] directions, respectively; ▲ and ×: respective values calculated using Eq. (1); ★: values computed using our program.

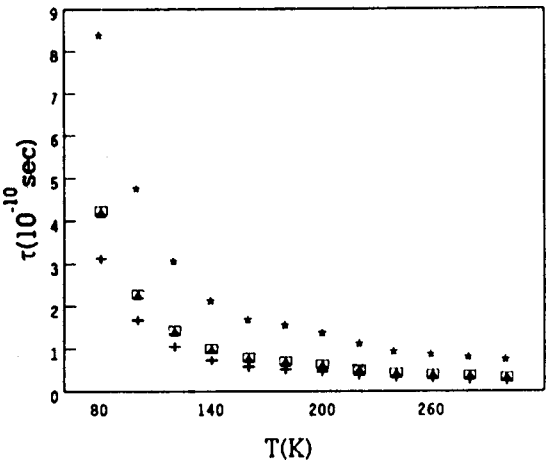


FIG. 2. Same as Fig. 1 for GaAs. +, □, ▲: values calculated using Eq. (1) for longitudinal waves along [100], [110], and [111] directions, respectively; ★: values computed using our program.

dence of relaxation time for longitudinal waves along the [100] and [110] directions. This figure also includes the relaxation times along these directions calculated using Eq. (1). The calculation of τ along different directions is not facilitated with our present computational scheme.

It can be seen from Fig. 1 that the relaxation times calculated using our program show a better agreement with the experiment than those calculated using Eq. (1) over the temperature range under study. Though the trend of temperature dependence of relaxation time for both these directions is reproduced quite well by our calculations, and though there is a good quantitative agreement with experimental relaxation times for the [110] direction, our calculations show a significant departure from the experimental values for the [100] direction. Figure 2 shows results of our calculation of the temperature dependence of relaxation time for GaAs. For GaAs, comparison could not be made, as the experimental data on temperature dependence of relaxation times has not yet been reported.

The temperature dependence of various physical parameters required in the estimation of attenuation is presented in Table I. Calculated values of D and Γ^2 at room temperature are presented in Table II. The results of our calculation of temperature dependence of attenuation obtained using our program with the Merkulov refinement are presented for different directions in Figs. 3–5. In these figures, the results obtained using our refined calculation of relaxation time and those obtained using the thermal relaxation time calculated using Eq. (1) are compared with the available experimental data.

If we ignore the Merkulov correction factor in calculations based on Mason's theory, then the results are found to be almost the same as those obtained using Mason's pure-mode scheme, showing, however, a slight quantitative improvement in agreement with the experimental data (Figs. 1–6 of Ref. 27). We have observed this feature in the present work also, so results of Mason's scheme are not included in Figs. 3–5.

A glance at Figs. 3–5 shows that the results of calculation based on the correction factor suggested by Merkulov *et al.* in our program and the use of collective phonon-relaxation times obtained using our computational scheme are in closer agreement with the experimental attenuation than those based on the use of relaxation times obtained using Eq. (1) in all six cases studied. However, the results of

TABLE I. Temperature dependence of various parameters used in the calculation of attenuation for GaAs.

					V			
					(10^5cm/s)			
T (K)	α (10^{-6} K)	θ_D/T	C_v (10^7erg/cm^3 K)	E_0 (10^7erg/cm^3)	$V_1[110]$	$V_s[110][\bar{1}10]$	$V_s[110][001]$	$V_1[111]$
80	0.7899	4.3514	0.4159	11.4514	5.2695	2.4904	3.3622	5.4285
100	1.6190	3.4789	0.5369	21.0567	5.2663	2.4882	3.3599	5.4252
120	2.4570	2.8967	0.6250	32.7523	5.2623	2.4862	3.3572	5.4211
140	3.2334	2.4812	0.6883	45.9445	5.2588	2.4835	3.3548	5.4176
160	3.8644	2.1706	0.7341	60.1923	5.2557	2.4809	3.3533	5.4147
180	4.3324	1.9282	0.7682	75.2531	5.2515	2.4784	3.3510	5.4105
200	4.6857	1.7330	0.7940	90.9536	5.2463	2.4755	3.3468	5.4051
220	4.9680	1.5745	0.8136	107.0518	5.2425	2.4729	3.3445	5.4013
240	5.1928	1.4421	0.8290	123.5104	5.2388	2.4707	3.3420	5.3975
260	5.3668	1.3298	0.8412	140.2550	5.2343	2.4684	3.3389	5.3929
280	5.5025	1.2335	0.8511	157.2181	5.2293	2.4663	3.3359	5.3877
300	5.6190	1.1499	0.8590	174.3642	5.2243	2.4651	3.3328	5.3824

TABLE II. Mason's nonlinearity constant and Nava's ultrasonic Grüneisen parameter for different directions.

Direction of propagation	Direction of polarization	D	Γ^2
100	100	0.835	0.672
100	001	0.165	0.200
110	110	0.721	0.616
110	$1\bar{1}0$	0.623	0.244
110	001	0.169	0.200
111	111	0.683	0.599
111	$\bar{1}10$	1.604	0.229

our calculation using both these approaches for the [110] shear wave polarized along the [001] direction are not in good agreement with the experiment (Fig. 5).

We have also performed calculations of the temperature dependence of attenuation using Nava *et al.*'s formulation in our program. These results are compared with the experimental data in Figs. 6 and 7. These results again show good agreement with the experiment in all the cases studied. The comparison of these figures with Figs. 3–5 reveals that results obtained using Nava's scheme are in better accord with the experiment than those obtained using the Mason–Merkulov scheme. It can also be noted that the maximum difference between the experimental attenuation and the Mason–Merkulov calculation is larger than that for the calculation based on Nava's formulation in all these cases.

As the complete calculation of Γ^2 requires much more knowledge about the dispersion and anisotropy of the parameters τ, γ_{ij}, and β_i than is available, Nava *et al.*[8] have treated Γ^2 as a parameter whose magnitude and temperature dependence are to be determined from experimental data. However, agreement with the experimental data, obtained by us in all the cases under study, supports our present approach of using the temperature-dependent UGP.

It should be noted here that the reported experimental data for GaAs show considerable variation. The values of attenuation reported by Shukla *et al.*,[19] in particular, are much larger compared to those reported by other authors. The reasons for this discrepancy are not clear. This could be partly because of the fact that the samples studied are never exactly the same, and that a small difference in the nature of doping can result in large difference in their conductivities[50] and TOEC.[51] The results of our computation show quite a good agreement with the experimental data of Cottom *et al.*[24] for longitudinal waves along the [110] and [111] directions. However, our results are not in agreement with the data of Shukla *et al.* (not included here). Some of these problems are expected to be resolved if experiments to determine the relaxation time, thermal conductivity, TOEC, and attenuation are performed on the same samples. It should also be mentioned here that GaAs is different from other materials studied so far by us[27,33–37] in that its departure from isotropy,

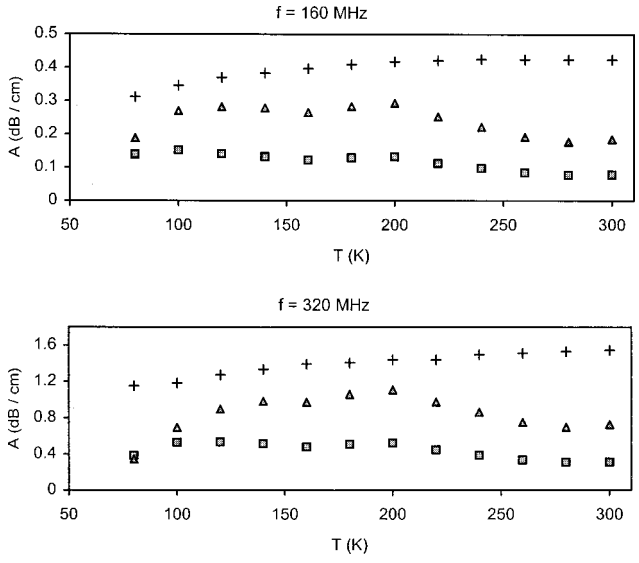


FIG. 3. Temperature dependence of ultrasonic attenuation for longitudinal waves along [110] direction. Values calculated using Mason–Merkulov scheme and ■: thermal τ of Eq. (1), ▲: our τ; +: experimental values from Ref. 24.

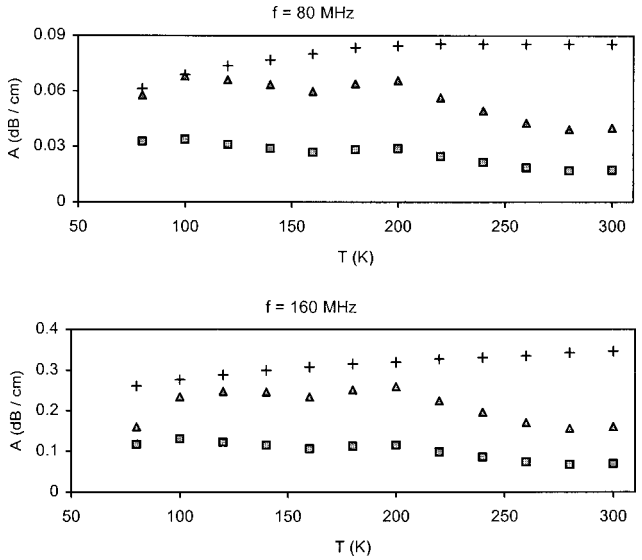


FIG. 4. Same as in Fig. 3 for longitudinal waves along [111] direction. Experimental values from Ref. 24.

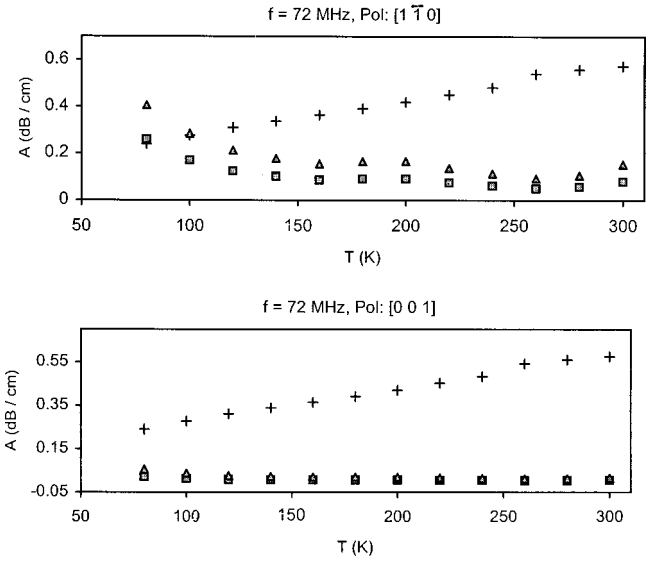


FIG. 5. Same as in Fig. 3 for shear waves along [110] direction. Experimental values from Ref. 23.

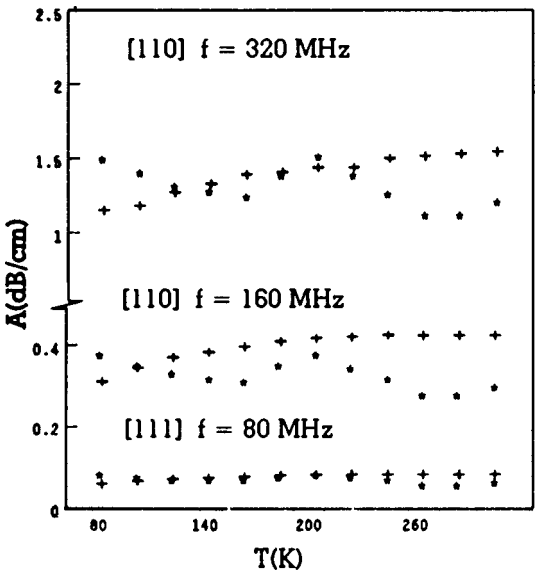


FIG. 6. Same as in Fig. 3 for longitudinal waves along [110] and [111] direction. ★: values calculated using Nava's formulation in our program, +: experimental values from Ref. 24.

as measured by the *s* and *t* parameters of de Launay,[45] is not regular with temperature. The anisotropy factors for TOEC defined by Brugger[52] also exhibit this anomaly in the temperature range under study.

Since the values measured by Shukla *et al.* were much larger than those obtained by those using Nava *et al.*'s formulation at room temperature, they determined the ratio of UGP at various temperatures by normalizing the UGP at 300 K for longitudinal waves and at 240 K for shear waves. Figure 8 shows the comparison of these ratios with our calculation. The agreement between the two is reasonable over the temperature range under study. The observations of Shukla *et al.* that the normalized UGP are the same for longitudinal waves along the [110] and [111] directions and for shear waves along the [100] and [111] directions are confirmed in the present study [Fig. 8(b) and (c)].

However, the calculated values of this ratio, as well as our calculated attenuation (Figs. 3–7), do not show a smooth trend, particularly in the high-temperature range. This can be attributed to the fact that the experimental TOEC values used in this calculation show significant variation and scatter in this temperature range.[26] The constants C_{112}, C_{123}, and C_{166}, in particular, vary considerably in this range.

We have also investigated the effect of incorporating our refined calculation of relaxation times in the calculation of attenuation using Nava *et al.*'s formulation. These results are in good agreement with the experimental trends but the attenuation is overestimated in most of the cases. As the exact calculation of attenuation using Nava *et al.*'s modified formulation of Woodruff's theory demands a rather complete knowledge of the dispersion and polarization dependence of the relaxation times and the Grüneisen parameters, incorporation of these factors in the present study is expected to shed more light on the present investigation. Finally, it should be mentioned that we have not investigated the anisotropy in relaxation times reported by Ilisavskii *et al.* for Si. Such an investigation demands inclusion of acoustic modes in the scheme of computation.

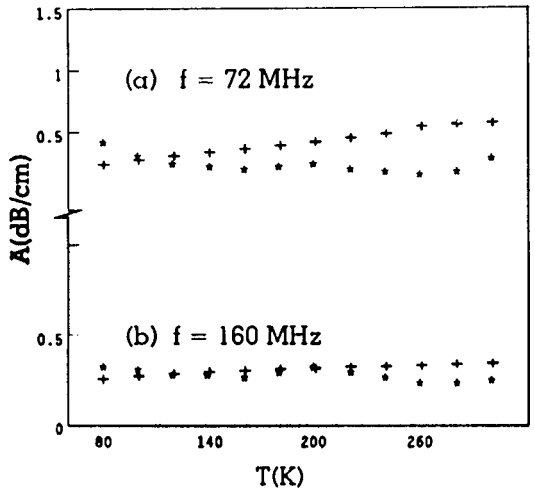


FIG. 7. Same as in Fig. 6 for (a) [110] shear wave polarized along $[1\bar{1}0]$ direction, and (b) longitudinal wave along [111] direction. Experimental values from Refs. 23 and 24.

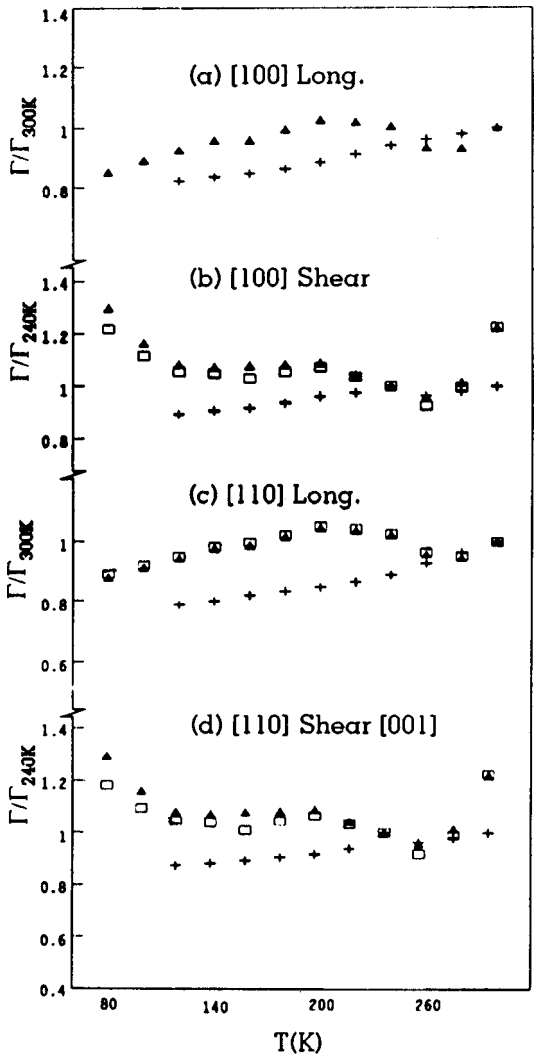


FIG. 8. Temperature dependence of the normalized UGP. ▲: Ratio calculated using our program; +: experimental values from Ref. 19; □: calculated ratio for (i) $[111][\bar{1}10]$ in (b), (ii) [111] long in (c) and (iii) $[110][1\bar{1}0]$ in (d).

[1] H. J. Maris, in *Physical Acoustics*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1971).
[2] L. Landau and G. Rumer, Phys. Z. Sowjetunion **11**, 18 (1937).
[3] A. Akhiezer, J. Phys. (Moscow) **1**, 277–287 (1939).
[4] T. O. Woodruff and H. Ehrenreich, Phys. Rev. **123**, 1553–1559 (1961).
[5] W. P. Mason, in *Physical Acoustics*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1964), Vol. IIIB, Chap. 6.
[6] H. H. Barrett and M. G. Holland, Phys. Rev. B **1**, 2538–2544 (1970).
[7] V. Ya. Avdonin, V. V. Lemanov, I. A. Smirnov, and V. V. Tikhonov, Sov. Phys. Solid State **14**, 747–752 (1972).
[8] R. Nava, M. P. Vecchi, J. Romero, and B. Fernandez, Phys. Rev. B **14**, 800–807 (1976).
[9] S. K. Kor and Kailash, J. Phys. Soc. Jpn. **55**, 2232–2234 (1986).
[10] S. K. Kor and R. K. Singh, Acustica **80**, 83–87 (1994).
[11] W. P. Mason and T. B. Bateman, J. Acoust. Soc. Am. **36**, 644–652 (1964).
[12] W. P. Mason and T. B. Bateman, J. Acoust. Soc. Am. **40**, 852–862 (1966).
[13] M. A. Breazeale and J. Philip, in *Physical Acoustics*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1984), Vol. XVII, Chap. 1.
[14] S. Rajagopalan and D. N. Joharapurkar, J. Appl. Phys. **54**, 3166–3171 (1983).
[15] S. Rajagopalan and D. N. Joharapurkar, J. Appl. Phys. **55**, 275–277 (1984).
[16] S. Rajagopalan and D. N. Joharapurkar, J. Appl. Phys. **56**, 1333–1337 (1984).
[17] D. N. Joharapurkar, J. Appl. Phys. **64**, 1726–1729 (1988).
[18] Yu. V. Ilisavskii and V. M. Sternin, Sov. Phys. Solid State **27**, 236–239 (1985).
[19] S. S. Shukla and S. S. Yun, J. Acoust. Soc. Am. **70**, 1713–1716 (1981).
[20] B. A. Bobylev and A. F. Kravchenko, Sov. Phys. Acoust. **12**, 315–317 (1967).
[21] L. Hrivnak and J. Kovar, Phys. Status Solidi **23**, 189–193 (1967).
[22] E. Groubert and A. M. Joullie, Phys. Status Solidi **39**, 357–362 (1970).
[23] Yu. Kh. Vekilov, A. E. Kadyshevich, and O. M. Krasilnikov, Sov. Phys. Solid State **13**, 1095–1102 (1971).
[24] R. I. Cottam and G. A. Saunders, J. Phys. C., Solid State Phys. **7**, 2447–2456 (1974).
[25] W. Chen and H. J. Maris, Philos. Mag. **70**, 687 (1994).
[26] D. N. Joharapurkar, D. Gerlich, and M. A. Breazeale, J. Appl. Phys. **72**, 2202–2208 (1990).
[27] S. D. Lambade, G. G. Sahasrabudhe, and S. Rajagopalan, Phys. Rev. B **51**, 15861–15866 (1995).
[28] M. F. Lewis, J. Acoust. Soc. Am. **44**, 713 (1968).
[29] L. G. Merkulov, R. V. Kovalenok, and E. V. Konovodehenko, Sov. Phys. Solid State **11**, 2241–2248 (1970).
[30] S. K. Kor, U. S. Tandon, and P. K. Mishra, J. Appl. Phys. **45**, 2396–2397 (1974).
[31] M. Nandanpawar and S. Rajagopalan, J. Acoust. Soc. Am. **71**, 1469–1472 (1982).
[32] R. Nava, J. Non-Cryst. Solids **76**, 413–418 (1985).
[33] S. D. Lambade, G. G. Sahasrabudhe, and S. Rajagopalan, Acoust. Lett. **18**, 146–150 (1995).
[34] G. G. Sahasrabudhe and S. D. Lambade, Communicated to Acustica (Germany) (1998).
[35] S. D. Lambade, G. G. Sahasrabudhe, and S. Rajagopalan, J. Phys. Chem. Solids **57**, 217–223 (1996).
[36] S. D. Lambade, G. G. Sahasrabudhe, and S. Rajagopalan, J. Appl. Phys. **78**, 6525–6533 (1995).
[37] G. G. Sahasrabudhe and S. D. Lambade, J. Acoust. Soc. Am. **104**, 81–85 (1998).
[38] M. G. Holland, Phys. Rev. **132**, 2461–2471 (1963).
[39] E. F. Steigmeier and I. Kudman, Phys. Rev. **141**, 767–774 (1966).
[40] R. A. H. Hamilton and J. E. Parrott, Phys. Rev. **178**, 1284–1292 (1969).
[41] Y. P. Joshi and G. S. Verma, Phys. Rev. B **1**, 750–755 (1970).
[42] K. Brugger, Phys. Rev. A **137**, 1826–1827 (1965).
[43] K. Brugger and T. C. Fritz, Phys. Rev. **157**, 524–531 (1967).
[44] H. M. Antia, *Numerical Methods for Scientists and Engineers* (Tata McGraw-Hill, New York, 1991), pp. 240–260.
[45] J. de Launay, J. Chem. Phys. **22**, 1676 (1954); **24**, 1071 (1956); **30**, 91 (1959).
[46] M. G. Holland, Phys. Rev. **134**, A471–480 (1964).
[47] *American Institute of Physics Handbook* (McGraw-Hill, New York, 1972), 3rd ed.
[48] R. I. Cottam and G. A. Saunders, J. Phys. C., Solid State Phys. **6**, 2105–2118 (1973).
[49] R. Truell, C. Elbaum, and B. Chick, *Ultrasonic Methods in Solid State Physics* (Academic, New York, 1969), pp. 14–15.
[50] M. G. Holland, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1966), Vol. 2, Chap. 1.
[51] J. R. Drabble and J. Fendley, Solid State Commun. **3**, 269–270 (1965).
[52] K. Brugger, J. Appl. Phys. **36**, 759–768 (1965).

Laser-generated nonlinear Rayleigh waves with shocks

A. Lomonosov and V. G. Mikhalevich
General Physics Institute, Russian Academy of Sciences, 117942 Moscow, Russia

P. Hess
Institute of Physical Chemistry, University of Heidelberg, D-69120 Heidelberg, Germany

E. Yu. Knight
Department of Physics, University of California—Berkeley, Berkeley, California 94720-7300

M. F. Hamilton and E. A. Zabolotskaya
Department of Mechanical Engineering, The University of Texas at Austin, Austin, Texas 78712-1063

(Received 16 January 1998; accepted for publication 18 December 1998)

Intense laser radiation was used to generate a Rayleigh wave pulse of finite amplitude in fused quartz. Measurements of the pulse at two locations along the propagation path reveal the formation of well-defined shocks in the horizontal (in-plane) velocity waveform. As the pulse propagates, the different propagation speeds of the head and tail shocks lead to a considerable increase in the duration of the pulse. Theoretical predictions based on nonlinear spectral evolution equations are in close agreement with the observed waveform distortion and shock formation. © *1999 Acoustical Society of America.* [S0001-4966(99)03704-2]

PACS numbers: 43.25.Fe [MAB]

INTRODUCTION

Rayleigh waveform distortion observed by Lomonosov and Hess[1] can be interpreted as a nonlinear evolution process culminating in the formation of well-defined shocks. Both generation and detection of the surface waves were performed with laser beams. The vertical (normal with respect to the interface) component of the particle velocity waveform measured at two different locations along the propagation path of a single Rayleigh wave pulse are presented, together with calculations of the corresponding horizontal (in-plane) component and frequency spectrum. The observed waveform evolution between the two detection points is compared here with theoretical predictions obtained from spectral evolution equations derived by Zabolotskaya.[2]

Laser irradiation is a common noncontact method for generating elastic waves in solids.[3] In nondestructive testing, the laser intensity is limited to the thermoelastic mode of operation, and the efficiency of elastic wave generation is typically so low that propagation of the resulting surface waves is adequately described by linear theory.[4] At higher laser intensities, at which optical breakdown occurs, surface waves have been generated with amplitudes so large that microparticles are ejected from the surface along the propagation path.[5]

The efficiency of surface wave generation in the thermoelastic mode of operation can be increased considerably by introducing on the surface of the solid a thin layer of liquid having a large optical absorption coefficient. In this method, absorption of the laser radiation produces intense evaporation of the liquid coating, and the resulting reaction pressure leads to generation of the surface wave. Generation of surface waves using a liquid coating has been shown to be substantially more efficient than generation by direct laser irradiation, in the absence of liquid coupling, even when optical breakdown occurs.[6] Laser-induced photoacoustic generation incorporating a liquid layer was the method used to produce the measured Rayleigh waves reported here.

Theories for nonlinear surface waves that were advanced in the 1980s[7–9] were motivated in part by the development of surface acoustic wave devices for nonlinear signal processing operations. The corresponding predictions of harmonic generation and waveform distortion were restricted to the preshock region. In contrast with these earlier models, the theory employed here to interpret measurements of nonlinear Rayleigh waves was derived using Hamiltonian mechanics.[2,10] Because the evolution equations follow directly from the Hamiltonian for nonlinear surface waves, the theory can be extended in a straightforward way to include effects such as anisotropy and piezoelectricity. This model has been used in theoretical investigations of waveform distortion[2] and harmonic generation[11] out to distances well beyond where shocks are formed.

We note that a similar comparison of theory and experiment was published recently by Kolomenskii *et al.*[12] In their work, waveforms were compared with calculations based on an evolution equation developed by Gusev *et al.*[13] One way their work differs from ours is that the material constants in their theory were adjusted to provide a best fit with the measurements. We do not employ curve fitting, and we use measurements of the third-order elastic constants that are published in the literature. Another difference is that the present theory describes the cusped wave profile near the shocks in the measured waveforms, a feature that has been discussed in detail previously[2,11] and which distinguishes the nonlinear distortion of surface waves from that of bulk waves.

I. EXPERIMENT

Our experiments investigated the propagation of nonlinear Rayleigh waves in fused quartz. The material is consid-

 0001-4966/99/105(4)/2093/4/$15.00 © 1999 Acoustical Society of America

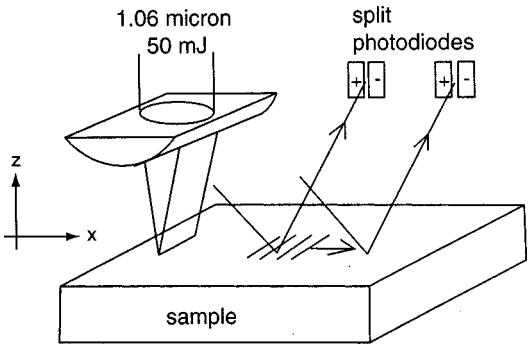


FIG. 1. Diagram and geometry of the experiment.

ered to be isotropic, and the small-signal elastic properties are characterized by the longitudinal wave speed $c_l = 5960$ m/s, transverse wave speed $c_t = 3754$ m/s, and density $\rho = 2203\,\mathrm{kg/m^3}$, for which the corresponding Rayleigh wave speed is $c_R = 3401$ m/s. Rayleigh wave pulses were generated by shining Nd:YAG laser radiation onto the material. The pulsed laser radiation was focused with a cylindrical lens to form a strip having dimensions 6 mm by 50 μm on the surface of the solid (see Fig. 1). The wavelength of the laser radiation was 1064 nm, the pulse length 7 ns, and optical pulse energies up to 100 mJ were used. At the location where the sample was irradiated by the laser, the surface of the solid was covered by a liquid layer having a large optical absorption coefficient.

The Rayleigh wave pulse generated by laser excitation was detected at two locations on the surface of the solid. The two locations were at distances 2.3 mm and 18.3 mm along a direction normal to the long (6 mm) dimension of the irradiated strip. A laser beam deflection technique was used to make the measurement at each location. A second Nd:YAG laser was used for this purpose, and the probe beams were formed by the second harmonic component of the cw laser radiation (532 nm, 30 mW). The spot size of each probe beam on the surface was 4 μm. Deflection of each beam was detected by a separate fast photodiode, each possessing 500-MHz bandwidth. The detected signal is proportional to the slope of the surface, and therefore, for a traveling surface wave, proportional to the vertical component of the surface velocity.

II. RESULTS

A vertical velocity waveform v_z measured at the first location, 2.3 mm from the source, is shown in Fig. 2(a). The corresponding horizontal velocity waveform v_x in Fig. 2(b) was calculated by taking the Hilbert transform of the vertical component:

$$v_x(t) = -\left(\frac{\eta+\xi_t}{1+\eta\xi_l}\right)\frac{1}{\pi}\,\mathrm{Pr}\int_{-\infty}^{\infty}\frac{v_z(t')}{t'-t}\,dt', \tag{1}$$

where $\xi_l = (1-c_R^2/c_l^2)^{1/2}$, $\xi_t = (1-c_R^2/c_t^2)^{1/2}$, and $\eta = -2\xi_t/(1+\xi_t^2)$. We have taken x to be the coordinate along the direction of propagation, and z to be the coordinate

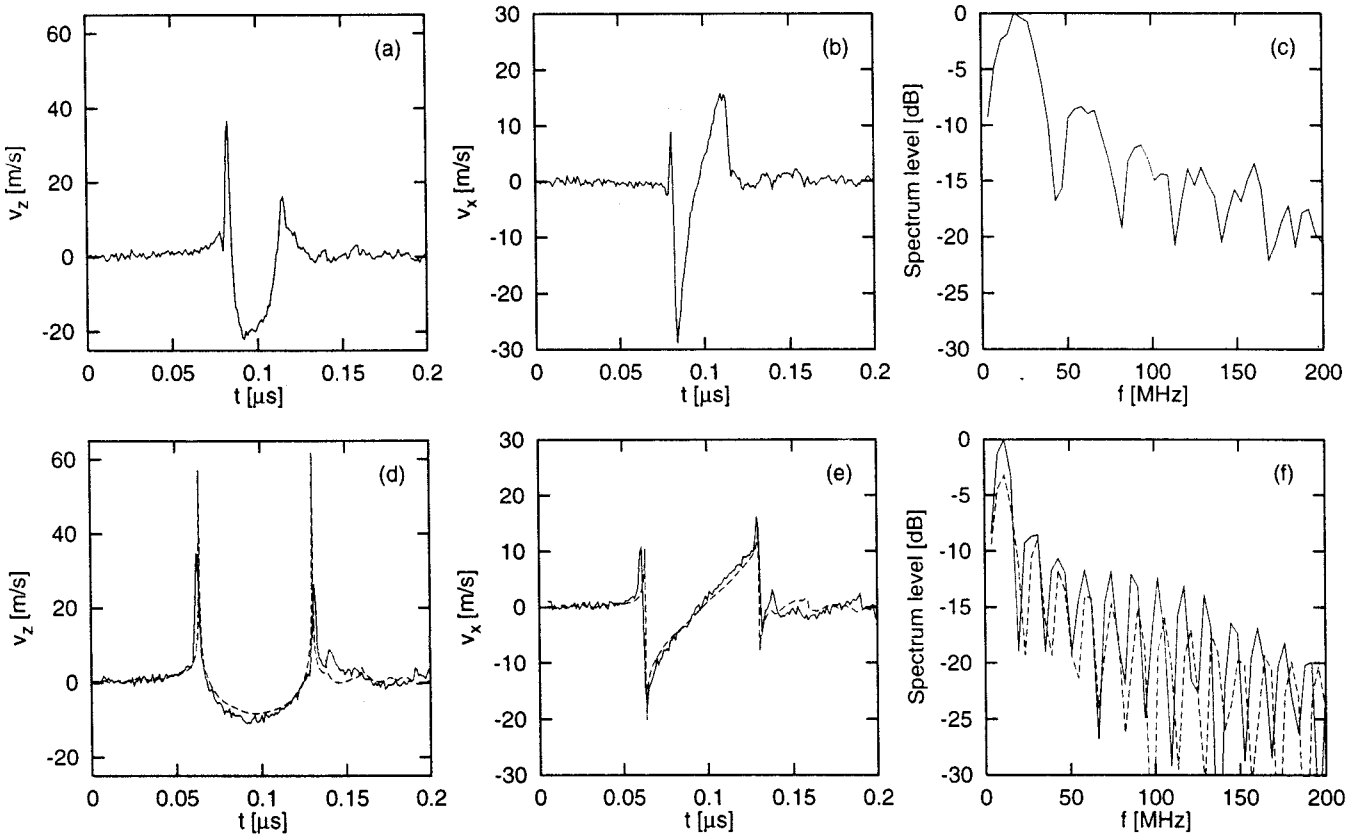


FIG. 2. Rayleigh surface waveforms and corresponding frequency spectra at distances 2.3 mm [upper row, (a)–(c)] and 18.3 mm [lower row, (d)–(f)] from the laser excitation. Solid lines are experiment, dashed lines are theory. The reference for the spectrum levels in (c) and (f) is the peak value of the measured spectrum at that location.

normal to, and defined positive outward from, the surface of the solid. The frequency spectrum of the signal, normalized by its peak value, is shown in Fig. 2(c).

We first assess the potential influence of diffraction on the propagation of the pulse. The characteristic diffraction length for a Rayleigh wave with angular frequency ω radiated from a line source with half-width a is $x_d = ka^2/2$, where $k = \omega/c_R$. Taking $a = 3$ mm and, from Fig. 2(c), $\omega/2\pi = 20$ MHz to characterize the dominant frequency band, yields $x_d = 166$ mm, which is an order of magnitude greater than the distance where the second measurement was made along the propagation path, $x = 18.3$ mm. Although the effect of diffraction was included in an earlier theoretical investigation of nonlinear Rayleigh wave beams,[14] its effect here is negligible and it is sufficient to use plane wave theory to interpret the observed waveform distortion.

The horizontal velocity waveform in Fig. 2(b) may be used to estimate the shock formation distance and thus the effect of nonlinearity over the propagation distance $\Delta x = 16$ mm between the two observation points. As discussed below, it is nonlinear distortion of the horizontal component, rather than the vertical component, that resembles the distortion of finite amplitude sound in fluids. The plane wave shock formation distance $\overline{x_R}$ for an initially sinusoidal Rayleigh wave in a lossless medium may be expressed in the same form as for sound waves in fluids:

$$\overline{x_R} = \frac{1}{|\beta_x| \epsilon_x k}, \tag{2}$$

where β_x is the coefficient of nonlinearity, and $\epsilon_x = v_{x0}/c_R$ is the peak acoustic Mach number (or equivalently, strain) associated with the initial waveform $v_x = v_{x0} \sin \omega t$. An approximate analytical expression for the coefficient of nonlinearity is given by Eq. (32) in Ref. 15. The value $\beta_x = -2.3$ is thus obtained for fused quartz from the values of the two second-order elastic constants, the bulk modulus $K = 36.9$ GPa and shear modulus $\mu = 31.0$ GPa (corresponding to the values cited above for c_l, c_t, and ρ), and the three third-order elastic constants measured by Bogardus.[16] Expressed in the notation of Landau and Lifshitz,[17] these third-order constants are $A = -42$ GPa, $B = 93$ GPa, and $C = 26$ GPa. Letting $v_{x0} = 30$ m/s and $\omega/2\pi = 20$ MHz characterize the waveform in Fig. 2(b), one obtains $\overline{x_R} = 1.3$ mm from Eq. (2).

The predicted shock formation distance is thus an order of magnitude less than the propagation distance $\Delta x = 16$ mm between the two locations where the waveforms were detected, which is consistent with the strong nonlinear distortion observed at the second location, shown as solid lines in Fig. 2(d) and (e). The negative value obtained for β_x indicates that nonlinearity causes positive portions of the horizontal velocity waveform to propagate with local phase speeds that are slower than the small-signal speed c_R of zero crossings, and negative portions to propagate faster, also in agreement with the distortion process observed by comparing Fig. 2(b) and (e). Because the negative head shock in the horizontal velocity waveform propagates faster than the positive tail shock, the pulse duration approximately doubles between the two observation points, which accounts for the corresponding spectral shift observed by comparing Fig. 2(c) and (f). Such distortion is opposite that of sound waves in fluids in the sense that positive portions of the acoustical particle velocity steepen forward and thus advance in time, whereas negative portions recede. However, the value of β_x may be either positive or negative for Rayleigh waves, and for most of the materials considered previously[15] it is positive.

The theoretical predictions shown as dashed lines in Fig. 2(d)–(f) were obtained by assuming plane wave propagation between the two observation points and solving the following set of coupled spectral equations[2] numerically:

$$\frac{dv_n}{dx} + \alpha_n v_n = -\frac{\mu \omega_0 n^2}{2\rho c_R^4 \zeta} \sum_{l+m=n} \frac{lm}{|lm|} R_{lm} v_l v_m, \tag{3}$$

where $v_n(x)$ is the amplitude of the nth harmonic component, ω_0 the angular repetition frequency for periodic signals (corresponding to the fundamental component $n = 1$), ζ a combination of second-order elastic constants, and R_{lm} a nonlinearity matrix that depends on both second- and third-order elastic constants.[11] The absorption terms $\alpha_n v_n$ were introduced *ad hoc* to ensure numerical stability when shocks develop in the waveforms. The horizontal and vertical velocity waveforms at the surface of the solid are constructed from the complex spectral amplitudes as follows:

$$v_x(x,t) = (\eta + \xi_t) \frac{i}{2} \sum_{n=-\infty}^{\infty} \frac{n}{|n|} v_n(x) e^{-in\omega_0(t - x/c_R)}, \tag{4}$$

$$v_z(x,t) = (1 + \eta \xi_l) \frac{1}{2} \sum_{n=-\infty}^{\infty} v_n(x) e^{-in\omega_0(t - x/c_R)}. \tag{5}$$

Equations (4) and (5) are related by Eq. (1).

Equations (3) were integrated with a fourth-order Runge–Kutta routine to simulate propagation of the signal measured at $x = 2.3$ mm out to the second observation point, $x = 18.3$ mm.[18] The pulse at $x = 2.3$ mm was assumed to repeat periodically every $T_0 = 0.255$ μs, and therefore the fundamental frequency in the Fourier series expansion is $\omega_0/2\pi = 3.92$ MHz. The computations were performed with $N = 800$ harmonics (i.e., $1 \leqslant n \leqslant N$, with $v_{-n} = v_n^*$). Initial values for v_n were determined via spectral decomposition of the measured signal shown in Fig. 2(a).

The absorption coefficients were chosen by assuming classical energy dissipation due to viscosity and heat conduction, for which the quadratic frequency dependence $\alpha_n = n^2 \alpha_1$ is obtained.[19] Numerical values were assigned to the absorption coefficients by selecting $1/\alpha_m = 500\overline{x_R}$, where m designates the value of n for the harmonic component having a frequency characteristic of the initial waveform (here $m = 6$, corresponding to $\omega/2\pi = 23.5$ MHz). With the absorption length chosen to be substantially larger than both the shock formation distance and the total propagation distance, the effect of absorption is assumed to be very weak in comparison with that of nonlinearity, and its main influence is on the rise times of the shocks. All remaining coefficients, the

parameter ζ and matrix elements R_{lm}, were calculated directly from values of the fundamental material constants (ρ,μ,K,A,B,C).

Comparison of the dashed and solid lines in Fig. 2(d) and (e) reveals close agreement between theory and experiment. Consider first the horizontal velocity waveforms in Fig. 2(e). The appearance of cusped spikes at the shocks, which was predicted to be a distinctive characteristic of nonlinear Rayleigh waves,[2] is clearly manifest in the measured waveform, and the calculations accurately reproduce this feature. The cusping at the shocks in the horizontal velocity waveform is associated with nonlocal nonlinear effects,[20] which are absent from the exclusively local nonlinear distortion process for sound waves in fluids. Nonlinearity causes sound waves to develop sawtoothlike profiles devoid of spikes.[21] Measured and predicted vertical velocity waveforms are compared in Fig. 2(d), which exhibit impulses where shocks appear in the horizontal velocity waveforms. Discrepancies between the measured and predicted amplitudes of the spikes are due to experimental and numerical limitations.

Calculations were also performed with the third-order constants measured by Yost and Breazeale[22] for fused quartz. Their values yield $\beta_x=-2.1$, and therefore waveform distortion is predicted to occur approximately 10% more slowly than on the basis of the constants measured by Bogardus[16] and used in the calculations for Fig. 2. Apart from this difference in nonlinear length scales, the waveforms calculated with both sets of constants are virtually identical.

III. SUMMARY

Close quantitative agreement with measurements of shock formation in nonlinear Rayleigh waves was achieved with predictions based on Zabolotskaya's theoretical model.[2,10] Not only does the theory predict the rate of nonlinear evolution, but also the fine structure associated with cusping of the waveform near the shocks, a feature that distinguishes nonlinear distortion of surface waves from that of bulk compressional and shear waves. No curve fitting was employed, with all necessary third-order elastic constants taken from the literature.

ACKNOWLEDGMENTS

Yu. A. Il'inskii is thanked for helpful discussions of this work. The experiments were performed while one of the authors (A.L.) was on leave at University of Heidelberg with financial support from Volkswagen-Stiftung, the International Science Foundation, and the Russian Foundation for Basic Research. The authors on the U.S. side were supported by the National Science Foundation and the Office of Naval Research.

[1] A. Lomonosov and P. Hess, ''Laser excitation and propagation of nonlinear surface acoustic wave pulses,'' in *Nonlinear Acoustics in Perspective*, edited by R. J. Wei (Nanjing University Press, Nanjing, 1996), pp. 106–111.

[2] E. A. Zabolotskaya, ''Nonlinear propagation of plane and circular Rayleigh waves in isotropic solids,'' J. Acoust. Soc. Am. **91**, 2569–2575 (1992).

[3] C. B. Scruby and L. E. Drain, *Laser Ultrasonics* (Adam Hilger, Bristol, 1990).

[4] Al. A. Kolomenskii and A. A. Maznev, ''Surface responses in the laser irradiation of a solid: Rayleigh waves and precursors,'' Sov. Phys. Acoust. **36**, 258–261 (1990).

[5] A. A. Kolomenskii and A. A. Maznev, ''Shake-off of mechanical microparticles from a silicon surface with surface acoustic waves induced by laser pulses,'' Sov. Tech. Phys. Lett. **17**, 483–484 (1991).

[6] Al. A. Kolomenskii, A. A. Maznev, and V. G. Mikhalevich, ''Laser optoacoustic effect at the boundary of a strongly absorbing liquid and applications of this effect,'' Izv. Akad. Nauk SSSR, Ser. Fiz. **54**, 2451–2457 (1990).

[7] N. Kalyanasundaram, ''Nonlinear surface acoustic waves on an isotropic solid,'' Int. J. Eng. Sci. **19**, 279–286 (1981).

[8] N. Kalyanasundaram, R. Ravindran, and P. Prasad, ''Coupled amplitude theory of nonlinear surface acoustic waves,'' J. Acoust. Soc. Am. **72**, 488–493 (1982).

[9] D. F. Parker, ''Waveform evolution for nonlinear surface acoustic waves,'' Int. J. Eng. Sci. **26**, 59–75 (1988).

[10] E. Yu. Knight, M. F. Hamilton, Yu. A. Il'inskii, and E. A. Zabolotskaya, ''General theory for the spectral evolution of nonlinear Rayleigh waves,'' J. Acoust. Soc. Am. **102**, 1402–1417 (1997).

[11] D. J. Shull, M. F. Hamilton, Yu. A. Il'insky, and E. A. Zabolotskaya, ''Harmonic generation in plane and cylindrical nonlinear Rayleigh waves,'' J. Acoust. Soc. Am. **94**, 418–427 (1993).

[12] Al. A. Kolomenskii, A. M. Lomonosov, R. Kuschnereit, P. Hess, and V. E. Gusev, ''Laser generation and detection of strongly nonlinear elastic surface pulses,'' Phys. Rev. Lett. **79**, 1325–1328 (1997).

[13] V. E. Gusev, W. Lauriks, and J. Thoen, ''Theory for the time evolution of nonlinear Rayleigh waves in an isotropic solid,'' Phys. Rev. B **55**, 9344–9347 (1997).

[14] D. J. Shull, E. E. Kim, M. F. Hamilton, and E. A. Zabolotskaya, ''Diffraction effects in nonlinear Rayleigh wave beams,'' J. Acoust. Soc. Am. **97**, 2126–2137 (1995).

[15] E. Yu. Knight, M. F. Hamilton, Yu. A. Il'inskii, and E. A. Zabolotskaya, ''On Rayleigh wave nonlinearity, and analytical approximation of the shock formation distance,'' J. Acoust. Soc. Am. **102**, 2529–2535 (1997).

[16] E. H. Bogardus, ''Third-order elastic constants of Ge, MgO, and fused SiO_2,'' J. Appl. Phys. **36**, 2504–2513 (1965).

[17] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity*, 3rd ed. (Pergamon, New York, 1986), p. 107.

[18] E. Yu. Knight, ''Generalization of the theory for nonlinear Rayleigh waves to nonplanar and transient waveforms, and investigation of pulse propagation,'' M.A. thesis (The University of Texas at Austin, 1995).

[19] R. W. Lardner, ''Nonlinear Rayleigh waves: Harmonic generation, parametric amplification, and thermoviscous damping,'' J. Appl. Phys. **55**, 3251–3260 (1984).

[20] M. F. Hamilton, Yu. A. Il'insky, and E. A. Zabolotskaya, ''Local and nonlocal nonlinearity in Rayleigh waves,'' J. Acoust. Soc. Am. **97**, 882–890 (1995).

[21] D. T. Blackstock, M. F. Hamilton, and A. D. Pierce, ''Progressive waves in lossless and lossy fluids,'' in *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, Boston, 1998), Chap. 4.

[22] W. T. Yost and M. A. Breazeale, ''Adiabatic third-order elastic constants of fused silica,'' J. Appl. Phys. **44**, 1909–1920 (1973).

Outer-flow effects on turbulent boundary layer wall pressure fluctuations

Timothy A. Brungart, Gerald C. Lauchle,[a] Steven Deutsch, and Eric T. Riggs
Applied Research Laboratory, The Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804-0030

(Received 12 November 1997; revised 4 November 1998; accepted 16 January 1999)

The outer-flow contribution to the pressure fluctuations occurring at the wall beneath a turbulent boundary layer was studied experimentally. A moving wall wind-tunnel facility was developed for the work. A flat test plate was suspended at various heights over the movable tunnel wall such that interacting and noninteracting turbulent boundary layers were developed in the resultant channel. Mean and fluctuating velocity components were measured for cases with and without wall motion. Pressure fluctuations were measured, with pinhole microphones on the surface of the test plate forming the upper-channel boundary, at corresponding test conditions. The data show that the wall-pressure fluctuations are relatively insensitive to the details of the outer flow, even over the range of frequencies dominated by outer-flow turbulence structures. © *1999 Acoustical Society of America.* [S0001-4966(99)02904-5]

PACS numbers: 43.28.Ra, 43.30.Nb, 43.50.Nm [LCS]

INTRODUCTION

Wall-pressure fluctuations induced by turbulent flow have been studied extensively over the past four decades both theoretically[1–6] and experimentally.[7–12] Theoretical work has focused on understanding how turbulent velocity fluctuations induce wall-pressure fluctuations. Experimental efforts have focused on obtaining spectral measurements of the wall-pressure field free from background noise and transducer spatial resolution limitations. The work has been motivated by the need to reduce noise levels in engineering applications ranging from sonar systems to automobiles to aircraft.

In spite of the vast quantity of work performed to date, significant issues still remain. Keith *et al.*[13] note that for the data sets they examined, pipe flow wall-pressure spectra, $G_T(\omega)$, exhibit lower levels than comparable external flows when scaled on outer-flow variables: $[G_T(\omega)/\rho^2\delta^* U_e^3]$ vs $(\omega\delta^*/U_e)$. Here, ρ is the fluid density, δ^* is the boundary layer displacement thickness, and U_e is the free-stream velocity. They support this observation with speculation that spectra measured in a developing flow should exhibit some of the characteristics of both internal and external flows. Here, internal refers to those flows that are completely dominated by viscous effects such as fully developed pipe and channel flows. External refers to boundary layer flows where viscous effects are confined to a region very near a bounding surface. Studies of developing channel flow[13,14] show that, over the range $0.1<\omega\delta^*/U_e<0.3$, as the flow develops, the spectra decrease from the levels of the external flows to the levels of the internal pipe flows. Others[11,13] point out that the physical features of the outer flow for the internal and external flow cases are distinctly different and that these differences may account for the differences observed in the wall-pressure fluctuations. Measurements of the wall-pressure fluctuations on the fuselage of a sailplane,[15] and in a wind tunnel,[16] indicate the spectral levels to be well correlated with velocity fluctuations at a distance of 2.6 δ above the wall, where δ is the boundary layer thickness. These velocity fluctuations result from the large-scale irrotational motion above the boundary layer and are absent in internal flows.

In order to provide additional insight into the contribution of the outer flow to the wall-pressure fluctuations, experiments were conducted where the outer flow of an equilibrium turbulent boundary layer was systematically modified and the effects on the wall-pressure fluctuation statistics were examined. An equilibrium boundary layer was developed in a channel. It was modified by (1) varying the height of the upper channel wall, and (2) moving the lower boundary at the same speed and direction as that of the mean flow.

I. EXPERIMENTAL APPARATUS AND MEASUREMENT TECHNIQUES

A. Moving wall wind tunnel

A moving wall wind-tunnel facility[17] was used for this investigation, and is shown schematically in Fig. 1. A muffler is soft coupled to the blower inlet and the outlet air passes through a rubber bellows to an acoustically treated labyrinth inside the blower box. The flow is diffused in the wind tunnel over a distance of 0.819 m to a square settling section measuring 1.473 m on a side. The diffuser consists of four screened sections increasing monotonically in area with streamwise distance. The settling section cross-sectional area is constant, has a length of 0.619 m, and is lined with seven fine screens designed to reduce both free-stream velocity nonuniformities and turbulence levels. Downstream from the settling section is a 64:1, 1.16-m long contraction nozzle which discharges into a 0.294-m diameter, semicylindrical test section.

The moving wall wind tunnel was placed inside the

[a]Also with the Graduate Program in Acoustics.

 0001-4966/99/105(4)/2097/10/$15.00 © 1999 Acoustical Society of America

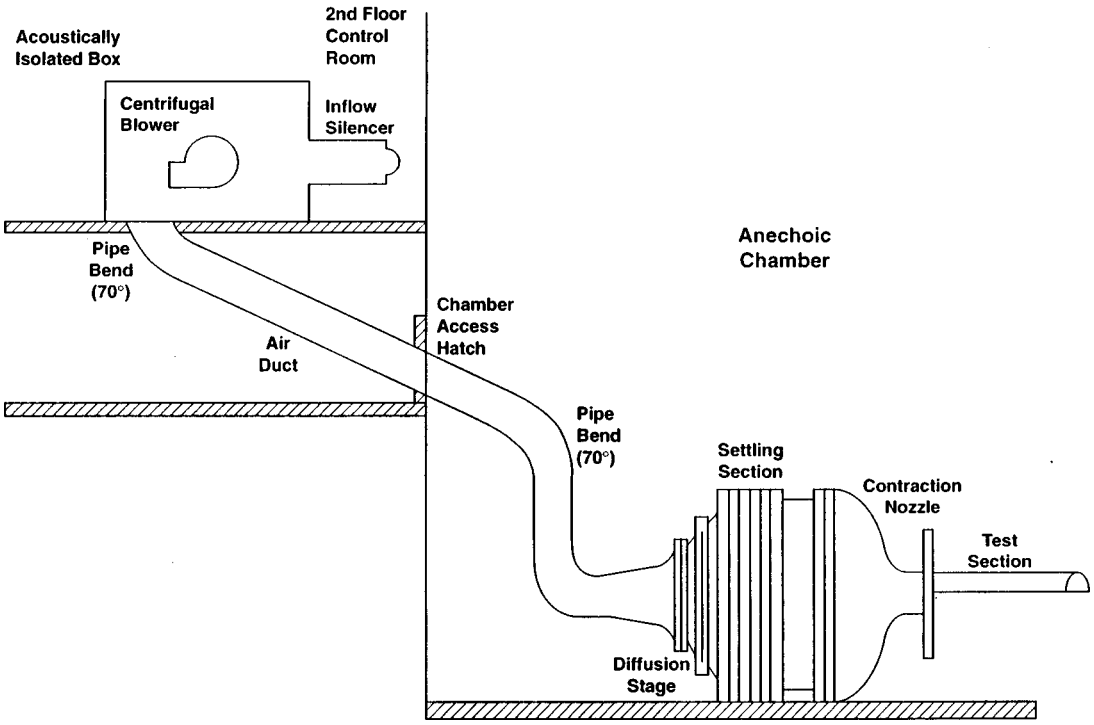


FIG. 1. Schematic of the moving-wall wind-tunnel facility.

ARL Penn State flow-through anechoic chamber[18,19] in order to isolate the tunnel from airborne noise. This chamber meets the requirements of both ISO 3745 and IEC 268 from 90 Hz to 12.5 kHz. Sound-pressure measurements can be made to ISO tolerances within 1.0 to 4.0 m of a source at frequencies as low as 70 Hz. The moving air generated by the wind tunnel passes through a permeable wall of the chamber with no measurable (above ambient) noise.

The moving wall wind-tunnel test section is shown schematically in Fig. 2. Slots are cut into the Plexiglas dome to permit measurement probe access inside the test section. The moving-wall assembly is vibration isolated from the test section. The moving wall is powered by an electric motor that is located in an acoustically isolated enclosure placed adjacent to the wind-tunnel test section. The motor drives the moving wall at a maximum speed of 18.3 m/s. The wall is a 0.019-cm thick Mylar belt which slides over Teflon pads. The drive roller is connected to the motor shaft with a flexible coupling. A fully developed turbulent boundary layer profile is generated only a few centimeters downstream from the entrance to the test section by lining the joint section with 36-grit sandpaper. This surface roughness trips the laminar boundary layer to a turbulent one.

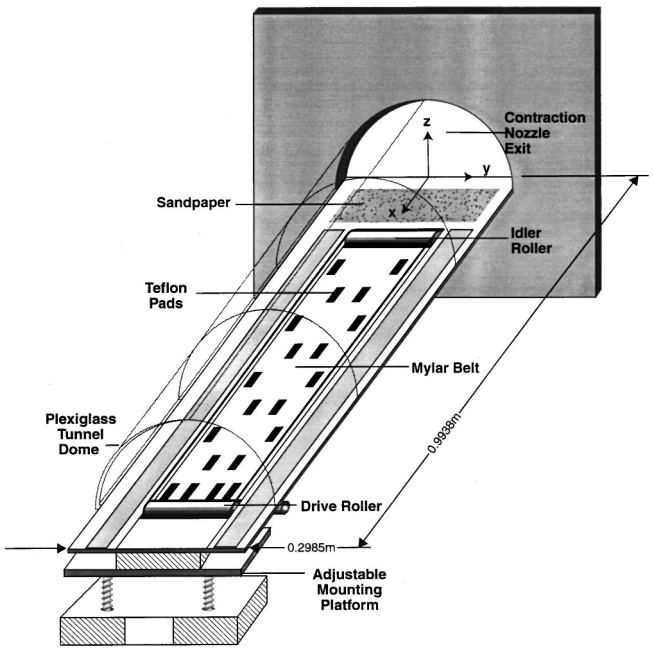


FIG. 2. Schematic of the moving-wall wind-tunnel test section.

The flat test plate is shown in Fig. 3. It is designed to span and fit tightly against the moving-wall test section at three heights above the moving surface. The leading edge was designed[20] to be resistant to flow separation due to small flow misalignments. The plate leading edge is positioned at the same axial location as the exit of the wind-tunnel contraction nozzle. Its leading edge is roughened with sandpaper to trip the boundary layer from laminar to turbulent.

Microphone measurement cavities are milled into the test plate in which small-diameter pinholes are used to connect the microphone cavity to the flow surface. This procedure helps to minimize spatial averaging effects on the plate turbulent boundary layer wall-pressure fluctuation measurements. The microphone is placed inside the cavity with its protective grid attached. The Helmholtz resonance frequency is estimated[21] to be 5.6 kHz. Figure 4 shows the pinhole microphone response to white-noise excitation in a plane wave tube. A peak in the response occurs near the predicted resonance frequency. The other peaks are acoustic resonances of the tube.

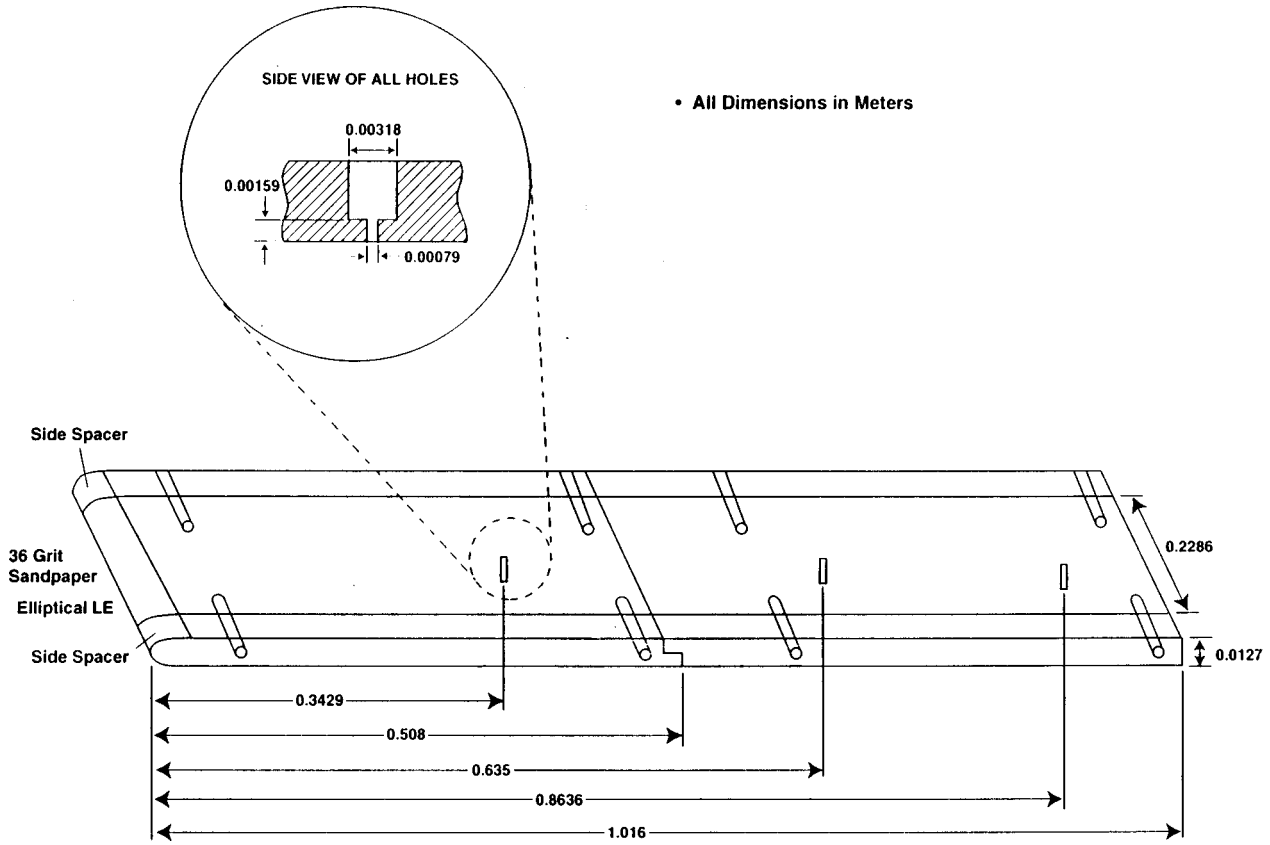


FIG. 3. Schematic of the test plate with pertinent dimensions.

B. Velocity measurements

For steady-state velocity measurements outside of the boundary layer, Pitot-static probes are used. For boundary layer velocity measurements, a flattened total head probe is used. A differential pressure transducer is used in conjunction with an integrating voltmeter to measure the pressures. The maximum uncertainty in the velocity measurements is estimated as $\pm 0.94\%$, $\pm 0.46\%$, and $\pm 0.24\%$ at flow speeds of 10.7, 15.3, and 21.4 m/s respectively.

A two-channel, TSI IFA 300 anemometer is used for measuring both mean and fluctuating velocities. Velocity statistics are calculated with TSI's THERMALPRO software. The hot-wire probes are aligned with the flow field and positioned with a traversing apparatus. A single-component TSI model 1218E standard probe and a TSI model 1241 end flow ''X'' probe are used. Hot-wire data are sampled at a frequency of 10 kHz with a low-pass antialiasing filter set at 5 kHz. Data are acquired for a period of 6.55 s providing 64 000 points at each measurement location. A typical outer-flow time scale is approximately 0.0013 s, so data are acquired over a period of 5000 time scales. A hot-wire overheat ratio of approximately 1.7 is used. The single-wire probes are calibrated *in situ* over a range of velocities while the X-wire probes are calibrated over a range of velocities and yaw angles in a special probe-calibration facility. Sensor yaw corrections are applied to the X-wire probes following the analysis of Lekakis *et al.*[22] The velocity calibrations are repeated prior to every measurement.

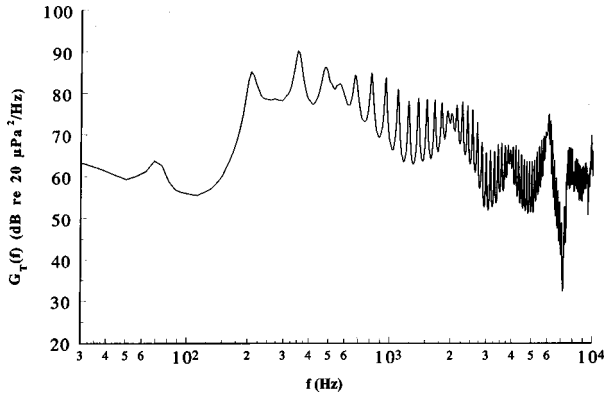


FIG. 4. Pinhole microphone response to white-noise excitation in a plane wave tube.

C. Wall-pressure fluctuation measurements

The turbulent boundary layer wall-pressure fluctuations are measured with a Brüel and Kjær (B&K) model 4138, 0.318-cm diameter condenser microphone placed in the seating hole of interest with a connecting preamplifier supported from above through the tunnel dome. The microphone signals are high-pass filtered above 20 Hz, amplified 10 to 20 dB, and processed with an FFT analyzer. Microphone calibrations were performed at the beginning of every test.

The turbulent boundary layer wall-pressure fluctuation signals are low-pass filtered at 10 kHz and sampled at 25.6 kHz with a record length of 4096 points. This provides an effective spectral bin width of 6.25 Hz. Typically, 512 and 256 averages are taken for each stationary and moving wall

condition, respectively, and a Hanning window function is used for all measurements. The normalized random error for a measured autospectrum, $\hat{G}_T(f)$, is given by[23] $\varepsilon_r[\hat{G}_T(f)] = (1/\sqrt{n_d})$ where n_d is the number of spectral averages. This gives values of 0.044 and 0.063 for 512 and 256 averages and corresponds to spectral estimates, with 95%-confidence intervals, of $\pm 8.8\%$ and $\pm 12.5\%$, respectively. Because the analysis bandwidth was not changed during the experiments, there are no bias errors.

II. RESULTS

A. Noninteracting boundary layers

This section presents the results of the flat plate flow and wall-pressure field evaluation where the plate and moving wall were separated by 5.40 cm, which is 2.4 plate boundary layer thicknesses based on measurements at X=0.883 m and a 15.3 m/s test velocity. This distance was sufficient to prevent the plate and Mylar belt shear layers from coalescing over the length of the moving-wall test section. The potential core of the exit wall jet is separated by the plate and Mylar belt shear layers throughout most of the test section length; thus, this flow will be referred to as external.

The streamwise variation in the free-stream velocity between the stationary belt forming the bottom test section wall and the flat plate at a reference speed of 16.9 m/s was found to be less than $\pm 1\%$. Operating the belt at the free-stream velocity tends to increase the free-stream velocity by approximately 1% at all measurement locations.

1. Wall boundary layers

The mean-velocity profiles on upper and lower walls were measured over the center 25% span with the belt held stationary, and found to be similar in the spanwise direction at all axial locations. All data shown were taken upstream from where three-dimensional corner disturbances contaminate the 2-D flowfield. Purtell *et al.*[24] state that examination of the boundary layer wake region is important not only to describe the boundary layer, but also in determining the state of development of the flow. All wall boundary layer mean-velocity profiles, measured at numerous streamwise (i.e., X) locations and free-stream speeds, were found to collapse on outer-flow variables (i.e., Z/δ^* vs U/U_e). The collapse of the data indicates that the boundary layer is fully developed at these conditions. See Brungart[17] for additional details.

The friction velocity, u_τ, and pressure-gradient parameter, Π, are determined by a least-squared error fit of the boundary layer velocity data to the law-of-the-wall and Coles'[25] wake function. A typical measured and predicted test plate boundary layer velocity profile, expressed in inner-flow variables (i.e., U/u_τ vs Zu_τ/v) is shown in Fig. 5. These results agree well and are typical of those obtained with the Mylar wall boundary layer as well. The measured Π value of 0.58 is consistent with the range of 0.5 to 0.6 identified by Coles[25] for a zero-pressure gradient flow. The data show further that, at the measurement location farthest downstream, X=0.883 m, the opposing boundary layers are still separated by 1.27 cm at a free-stream speed of 15.2 m/s.

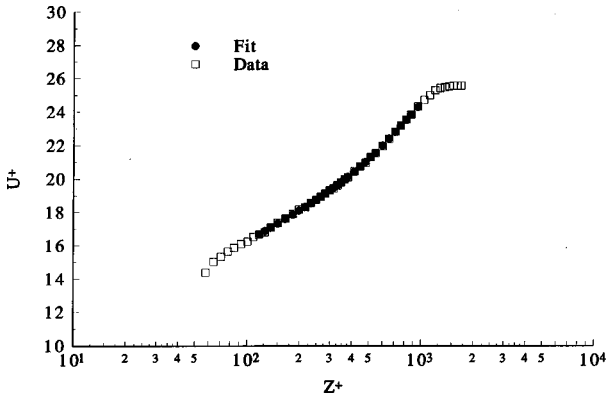


FIG. 5. Typical flat test plate boundary layer velocity profile expressed in inner-flow variables, wall stationary, X=0.883 m, $U_e=24.6$ m/s, $u^*=0.96$ m/s, $\Pi=0.58$.

Typical streamwise (i.e., u') and wall-normal (i.e., v') turbulent rms velocity fluctuation profiles for the two opposing walls, scaled on outer-flow variables, are shown in Fig. 6, and found to be in good agreement with data from Klebanoff.[26] The boundary layer displacement thickness, δ^*, used for nondimensionalization, is determined from integration of the measured mean-velocity profiles. The deviations above $Z/\delta^*=9$ are due to interaction effects with the Mylar wall boundary layer. The current turbulence stress data (not shown) are in good agreement with Klebanoff's[26] reported profile at values of Z/δ greater than 0.3. At values of Z/δ less than 0.3, the current data drop below Klebanoff's data and decrease as the wall is approached. This can be attributed to X-wire probe resolution limitations in the current experiments. The length scales in Klebanoff's flow were much larger than in the current experiments and probe resolution problems were not as severe.

Figure 7 compares the mean-velocity profile in the channel with and without wall motion. Positive Z is towards the moving wall. For the cases reported here, the wall is moving at a speed equal to the free-stream speed. Figure 7 shows that the boundary layer mean-velocity profile over the test plate is unmodified by the motion of the belt; however, the Mylar wall boundary layer mean-velocity profile becomes fuller with wall motion. The effect of wall motion on the streamwise turbulent rms velocity fluctuation component is shown

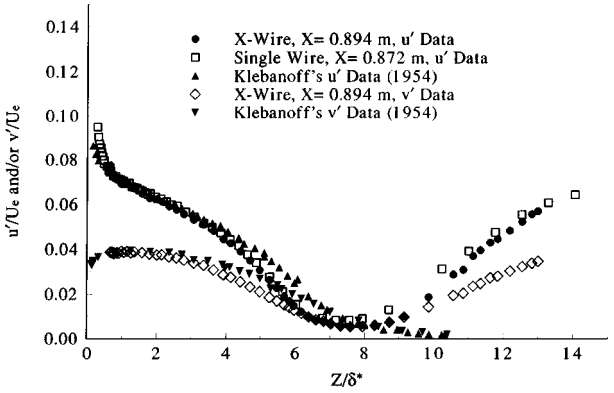


FIG. 6. Flat test plate boundary layer streamwise and wall-normal turbulence component rms velocity fluctuation profiles scaled with outer-flow variables and compared with data from Klebanoff (Ref. 26), $h_{ch}=5.40$ cm.

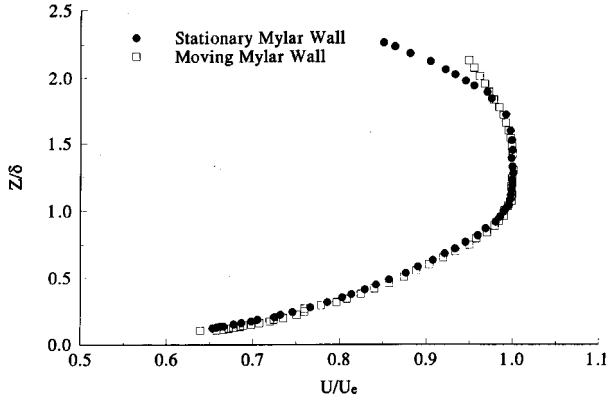


FIG. 7. Effect of wall motion on the flat test plate boundary layer velocity profile, X=0.894 m, U_e=15.4 m/s, h_{ch}=5.40 cm.

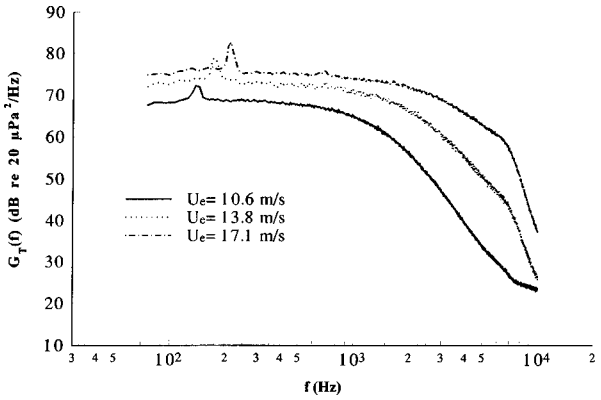


FIG. 9. Typical flat test plate boundary layer pressure fluctuation spectra, wall stationary, X=0.864 m, h_{ch}=5.40 cm.

in Fig. 8. The data show that the only effect of the wall motion is to suppress the velocity fluctuations in the Mylar belt boundary layer. Similar results[17] were obtained for the v' and Reynolds stress profiles, $\rho\overline{u'v'}$. In this case of non-interacting boundary layers, wall motion reduces the magnitude of the production term in the turbulence kinetic energy budget, thereby suppressing the turbulent velocity fluctuations in the moving-wall layer only.

2. Wall-pressure fluctuations

Typical spectra of the plate turbulent boundary layer wall-pressure fluctuations are shown in Fig. 9. The spectra were measured at X=0.864 m (X/δ=42.1 at 10.6 m/s) at a number of free-stream speeds. The narrow bands of higher energy in the spectra at approximately 175 Hz at 13.8 m/s and 225 Hz at 17.1 m/s are aerodynamic noise associated with the test plate and are not present in an empty tunnel. Finite coherence values between two pinhole microphones separated by 0.229 m in the test plate confirmed the acoustic nature of this narrow-band energy. With this microphone separation, the turbulent boundary layer wall-pressure fluctuations are incoherent while acoustical signals are coherent within an acoustic wavelength.

The distinct ''breakpoint'' in the two highest speed spectra at approximately 6 to 7 kHz in Fig. 9, although higher than the 5.6 kHz determined from the acoustic data of Fig. 4, are due to the Helmholtz resonance excitation of the pinhole cavity/microphone system. Panton and Miller[27] studied excitation of flush-mounted Helmholtz resonators by turbulent boundary layer flow. They found that the boundary layer-induced Helmholtz frequency was, in general, slightly different from that produced by acoustic excitation because of the interaction of the turbulence with the acoustical (i.e., lumped air mass) motion in the orifice.

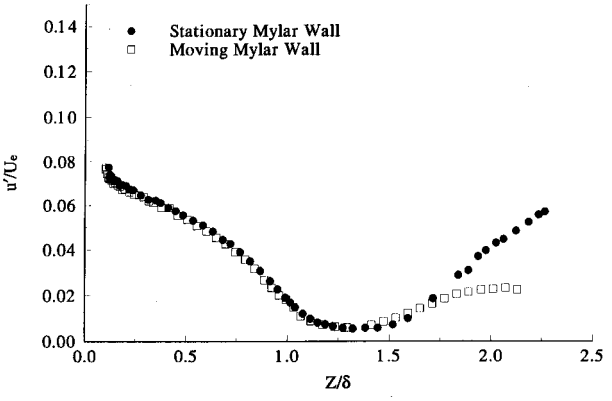


FIG. 8. Effect of wall motion on the flat test plate boundary layer streamwise turbulence component rms velocity fluctuation profile, X=0.894 m, U_e=15.4 m/s, h_{ch}=5.40 cm.

The steep, higher-frequency drop-off in the two higher-speed spectra in Fig. 9 is likely due to: (1) the Helmholtz resonance of the microphone/cavity system, and (2) spatial averaging of the wall-pressure fluctuations. In order to obtain a true point measurement of the turbulent boundary layer wall-pressure fluctuations, free from the effects of spatial averaging, the transducer diameter, $d_{\text{transducer}}$, must be of the same size or smaller than the smallest-length scale which characterizes the flow. The smallest-length scales present in a boundary layer, which give rise to measurable high-frequency pressure fluctuations, occur near the wall where the dynamics of the flow are influenced significantly by the effects of viscosity. A characteristic velocity fluctuation in this region is given by the friction velocity, u_τ, and the viscous length scale[28] is v/u_τ where v is the kinematic viscosity. Therefore, $d^+=(d_{\text{transducer}}u_\tau/v)$ must be $O(1)$ in order to resolve the high-frequency turbulent boundary layer wall-pressure fluctuations free from the effects of spatial averaging. Values of d^+ for the current measurements, at the location farthest downstream, range from 23.6 at 10.6 m/s to 45.1 at 13.8 m/s. The effect of spatial averaging is to reduce the microphone's sensitivity to high-wave number and high-frequency turbulence structures. According to analysis, based on the methods of Corcos,[29] the 10.6, 13.8, and 17.1 m/s spectra of Fig. 9 are attenuated by 3 dB at frequencies of 2830, 3600, and 4600 Hz, respectively, due to the effects of spatial averaging.

The moving wall introduces additional, unwanted tonal noise into the wind tunnel at frequencies below approximately 3 kHz. Comparison of pressure spectra measured with the wall moving and with and without air flow showed that a reasonable broadband signal-to-noise ratio exists in this frequency range. These comparisons also showed that at frequencies above approximately 3 kHz, the broadband noise

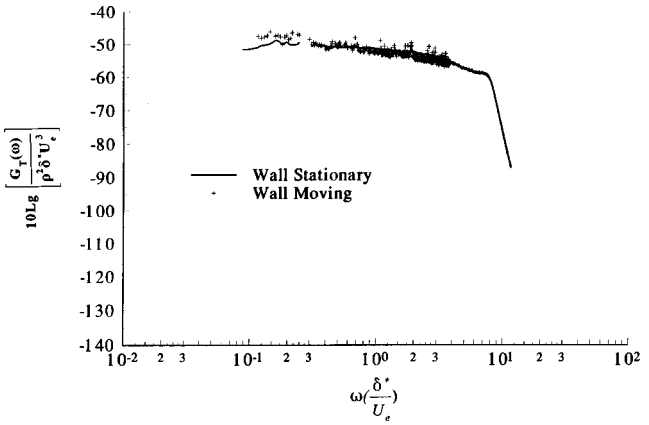


FIG. 10. Effect of wall motion on the flat test plate boundary layer pressure fluctuation spectrum, X=0.864 m, U_e=18.3 m/s, h_{ch}=5.40 cm.

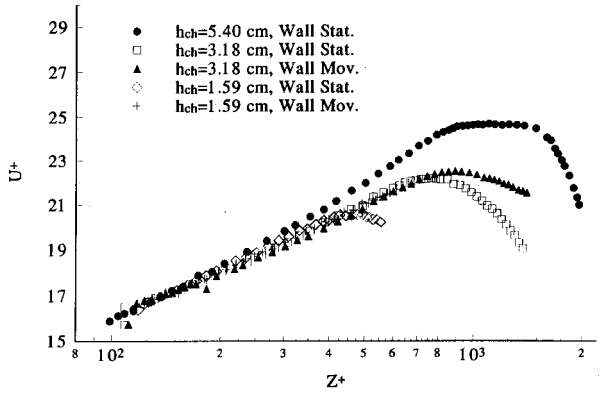


FIG. 11. Flat test plate boundary layer velocity profiles plotted in inner-flow variables for three channel heights, X=0.894 m, $U_{e/max}\approx$15.4 m/s.

associated with the belt sliding over the Teflon pads dominates the turbulent boundary layer wall-pressure fluctuation signal. Coherence measurements between the signals from the pinhole microphone and a frame-mounted accelerometer showed conclusively that the tones present in the spectrum of turbulent boundary layer wall-pressure fluctuations, when the belt is running, are a result of the vibration of the moving-wall apparatus.

A procedure[17] for removing the vibration-induced tonal signal from the wall-pressure fluctuation signal utilizes the measured coherence between the signal from the microphone and the signal from the accelerometer mounted on the frame of the moving-wall apparatus. Frequency ranges of high coherence ($\gamma^2 \geqslant 0.1$) were basically edited out of the pinhole microphone spectral data.

The effect of adjacent wall motion on the measured fixed-wall turbulent boundary layer wall-pressure fluctuations is shown in Fig. 10. These data, scaled on outer-flow variables, indicate that the wall motion has little if any effect on the measured wall-pressure fluctuations at this value of $h_{ch}/\delta = 2.75$, where h_{ch} denotes the channel height.

B. Developing flow results

This section presents the results of the flat plate flow and dynamic wall-pressure field evaluation for cases with and without wall motion where the test plate and Mylar wall boundary layers are interacting. The degree of the interaction is dependent on the axial location of interest and the distance separating the test plate from the Mylar wall. The plate/wall separation was reduced from $h_{ch}/\delta = 2.6$ used in Sec. A to $h_{ch}/\delta = 1.5$ and 0.7 for the current situation.

1. Boundary layer evaluation

The streamwise variation in the maximum core velocity measured between the stationary belt forming the bottom test-section wall and the plate mounted above it was evaluated by Pitot-static probe measurements. The measurements show that the flow decelerates somewhat through the channel. This effect is due, of course, to the coalescence of the opposing shear layers.

Test plate boundary layer mean-velocity profiles measured at X=0.318 m, for 10.8 m/s$\leqslant U_e \leqslant$18.3 m/s, and for the values of h_{ch}/δ noted above, were found to collapse well when normalized with outer-flow variables. This indicates that the boundary layer is fully developed. The impending downstream convergence of the shear layers was evident from the small spatial extent of uniform flow for the channel height of 1.5 δ. The shear layers coalesce at X=0.318 m when $h_{ch}/\delta = 0.7$. When normalized on inner-flow variables, the developing flow boundary layer mean-velocity profiles at this location are in excellent agreement with the logarithmic law-of-the-wall and Coles'[25] wake function. The measurements show that for $h_{ch} = 2.6\,\delta$ and 1.5 δ, the turbulent boundary layers are identical, fully developed, zero-pressure gradient, and under equilibrium conditions. The narrowest channel boundary layer, however, appears to be thinner and exhibits increased turbulence intensities relative to the larger plate/wall separations. This suggests that the interaction between the plate and Mylar wall shear layers has begun to modify the boundary layer flow upstream of the measurement location.

External and developing mean flow-velocity profiles measured farther downstream for both stationary and moving-wall conditions are shown in Fig. 11. The profiles are normalized with the test plate inner-flow variables and collapse well for $Z^+ < 200$. For $Z^+ > 200$, the values of U^+ for the largest channel height are greater than those of the developing flow cases. This indicates the formation of a well-defined outer flow.[25] The data for the narrower plate/wall separations follow the logarithmic law-of-the-wall (i.e., $\Pi = 0$) up to values of Z^+ near the channel centerline where the velocity reaches its maximum. This behavior is consistent with that of a fully developed turbulent channel flow as well as pipe-flow data.[30]

The rms value of the u' turbulence component, normalized by the local mean-flow velocity, U_{local}, is shown in Fig. 12. The wall-normal turbulence component and the turbulence stress profiles follow similar trends. The indicated behavior suggests that h_{ch} is not an appropriate scaling parameter for the external flow case. This makes physical sense because the plate boundary layer is unaffected by the presence of the Mylar wall spaced at $h_{ch} = 2.6\,\delta$.

Figure 13 shows that the rms values of the u' turbulence levels normalized by $U_{e/max}$ and expressed as a function of

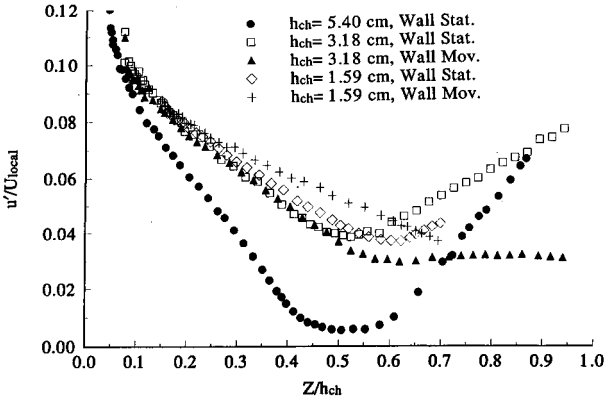


FIG. 12. Flat test plate boundary layer streamwise turbulence component rms velocity fluctuation profiles normalized with the local mean velocity and channel height, X=0.894 m, $U_{e/\max}\approx 15.4$ m/s.

Z/δ result in a slightly better collapse of the data than the normalization of Fig. 12. Here, $U_{e/\max}$ is the maximum mean velocity measured in the channel. The wall-normal turbulence component and the Reynolds stress data behave in a like manner. The plate external boundary layer can be viewed as having been modified by the presence of the Mylar wall boundary layer at values of Z greater than 0.7 δ. Identical outer-flow length scales characterize the plate boundary layer for $Z<0.7\,\delta$ in both the wider channels but not the narrow channel. This is reasonable, since the interaction between the plate and Mylar wall shear layers must be strongest for the narrow channel where the flow is most fully developed as a channel flow. The outer-flow length scale associated with the narrow channel must be smaller than that associated with the larger channels.

Figures 12 and 13 show that the rms turbulence fluctuations near the test plate ($Z\sim 0$) scale differently, depending on the channel width. This behavior indicates two flow types: one exhibiting the characteristics of an external turbulent boundary layer, and the other, a turbulent channel flow. The data for $h_{\rm ch}/\delta=1.4$ exhibits characteristics of each.

The effect of wall motion on the mean velocity is dependent on $h_{\rm ch}$. For $h_{\rm ch}/\delta=1.4$, moving the wall at the free-stream speed increases U_b, the semichannel bulk or average velocity, approximately 2%. For the case of $h_{\rm ch}/\delta=0.7$, U_b is increased 8%. The data show that the percent increase in u_τ due to wall motion is identical to the corresponding increase in U_b. For $h_{\rm ch}/\delta=1.4$, the effect of wall motion is limited to $Z/h_{\rm ch}>0.5$. Here, $u'/U_{e/\max}$ and $v'/U_{e/\max}$ are reduced by more than a factor of two (2) to near-constant values of 0.03 and 0.02, respectively. Similar trends are noted for the Reynolds stresses. The situation for $h_{\rm ch}/\delta=0.7$ is somewhat different because the changes occur closer to the wall.[17]

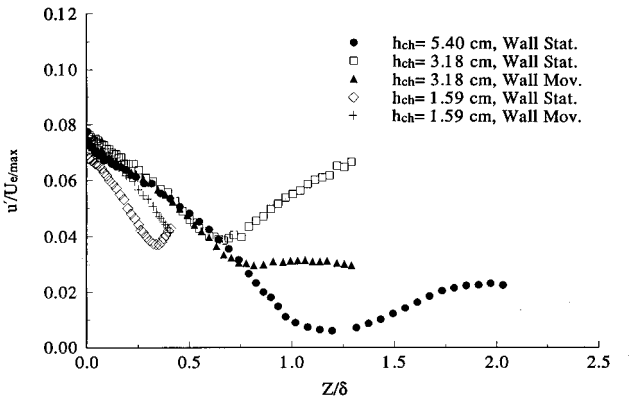


FIG. 13. Flat test plate boundary layer streamwise turbulence component rms velocity fluctuations normalized with the channel maximum velocity and expressed as a function of height above the surface, X=0.894 m, $U_{e/\max}\approx 15.4$ m/s.

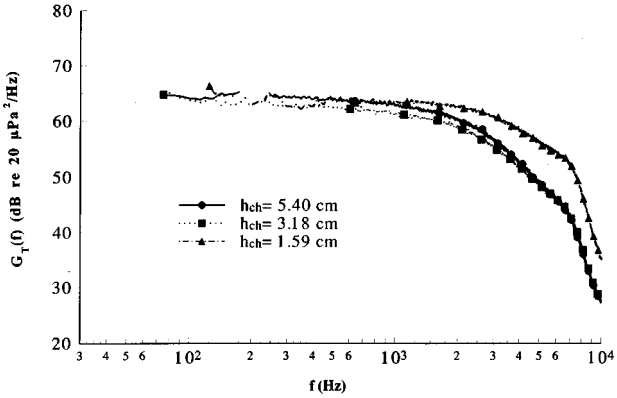


FIG. 14. Comparison of flat test plate boundary layer wall-pressure fluctuation spectra for three channel heights, wall stationary, X=0.864 m, $U_{e/\max}=15.3$ m/s.

2. Wall-pressure fluctuations

Wall-pressure fluctuation power spectra measured under stationary wall conditions are shown in Fig. 14 for the three plate/wall separations examined. The large and intermediate channel-separation spectra are almost identical. The plate-pressure fluctuations measured in the narrow channel situation exhibit increased high-frequency energy relative to the other separations. At low frequencies (up to 1200 Hz), the spectral levels of all separations considered are within 1 dB of each other.

Figure 15 compares wall-pressure fluctuation power spectra measured with and without wall motion for $h_{\rm ch}/\delta=1.4$. Clearly, wall motion has no significant effect on the stationary wall-point pressure spectrum. Figure 16 shows the

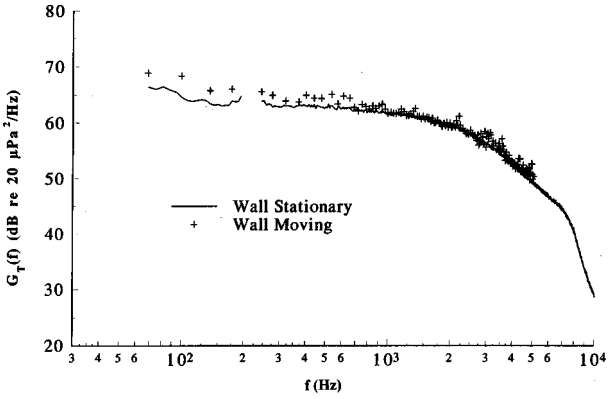


FIG. 15. Effect of wall motion on flat test plate boundary layer wall-pressure fluctuation spectra, X=0.864 m, $U_{e/\max}=15.9$ m/s, $h_{\rm ch}=3.18$ cm.

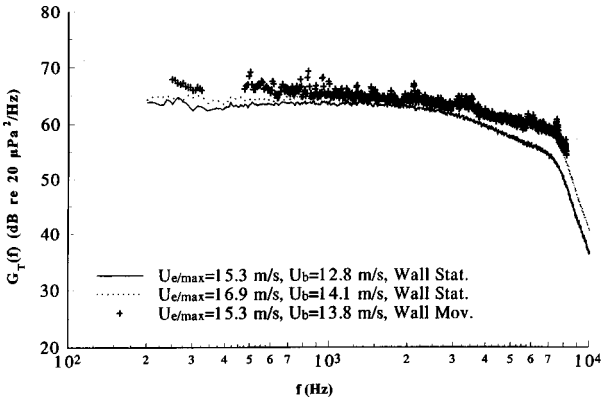


FIG. 16. Effect of wall motion on flat test plate boundary layer wall-pressure fluctuation spectra; comparison of spectra measured with approximately equivalent bulk velocity values, X=0.864 m, h_{ch}=1.59 cm.

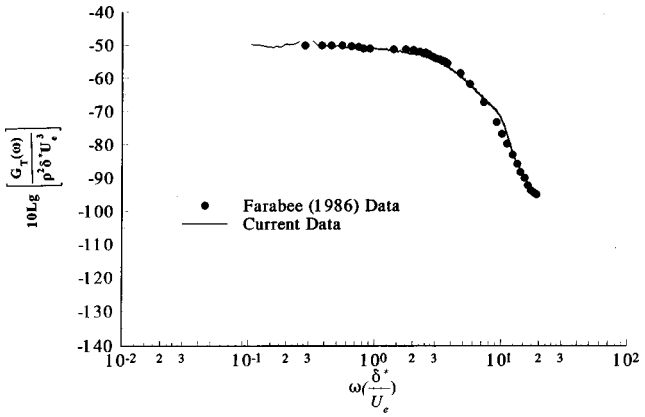


FIG. 17. Comparison of flat test plate boundary layer pressure spectrum with data from Farabee (Ref. 11), wall stationary, X=0.864 m, U_e =15.4 m/s, h_{ch}=5.40 cm.

effect of wall motion when the channel separation is decreased to 0.7 δ. We note that $U_{e/\max}$ is the moving-wall speed. The fixed wall-point wall-pressure spectrum under moving-wall conditions is higher in level than the corresponding stationary wall spectrum. However, it is the same as the stationary wall-pressure spectrum measured at a nearly equivalent value of U_b.

III. DISCUSSION

The current external flow turbulent boundary layer wall-pressure fluctuation measurements, nondimensionalized with outer-flow variables, are favorably compared to similar data measured by Farabee[11] in Fig. 17. Farabee's data were acquired with a pinhole microphone on the test-section wall of a low-noise wind tunnel with U_e=15.5 m/s, δ^*=0.450 cm, and d^+=33. The flow conditions for the subject data are U_e=15.3 m/s, δ^*=0.338 cm, and d^+=32.7.

Keith *et al.*[13] speculate that in developing flows, turbulent boundary layer wall-pressure spectra might display characteristics of both internal and external flows due to the existence of a free stream and the proximity of boundary layers on opposite and adjacent walls. In their[13] comparison of pressure spectra from numerous experimental investigations, developing channel flow measurements[14] in which the spectrum level decreases from the levels of external flows to the levels of internal pipe flows over the frequency range $0.1 < \omega\delta^*/U_e < 0.3$ support their argument. The channel height for these[14] measurements was 6 δ, and the span-wise width was 16 δ. The data shown in Fig. 17 are for 2.75 δ and 15 δ, respectively, yet the current measurements are almost identical to Farabee's[11] ''external'' flow measurements (i.e., test-section height and width both equal to 21.9 δ). This indicates that over the range of frequencies presented, the shear layers on adjacent and opposite tunnel walls have negligible effect on the measured turbulent boundary layer wall-pressure fluctuations.

Figure 18 shows an additional comparison of the current data at U_e=10.6 m/s, δ^*=0.366 cm, and d^+=23.7, with those of Schewe[10] at U_e=6.3 m/s, δ^*=0.460 cm, and d^+ =19. The two data sets are in excellent agreement and again show that the effect of the opposite and adjacent shear layers on the measured turbulent boundary layer wall-pressure fluctuations is negligible.

The moving wall-pressure spectral results for h_{ch} =2.6 δ (shown in Fig. 10) are consistent with the stationary-wall results. The velocity measurements, Figs. 7 and 8, showed that the wall motion influences only the belt boundary layer and has no effect on the plate turbulent boundary layer. The stationary wall data shows that boundary layers on opposite and adjacent tunnel walls have no effect on the wall-pressure fluctuations. Therefore, wall motion would not be expected to modify the wall-pressure fluctuations measured on the flat plate, and Fig. 10 shows this to be the case.

Farabee and Casarella[16] identify four regions in the turbulent boundary layer as dominating the wall-pressure fluctuations over four specific frequency ranges. The low-frequency range, defined by $\omega\delta/u_\tau \leqslant 5$, is dominated by sources associated with the large-scale structures induced by the unsteady potential flow observed above the boundary layer. Due to facility noise contamination, these sources could not be measured. A mid-frequency range is presumed dominated by structures in the outer region of the boundary layer (i.e., $5 \leqslant \omega\delta/u_\tau \leqslant 100$). It thus extends from the lowest frequencies measured to 528 Hz for the Fig. 14 external-flow

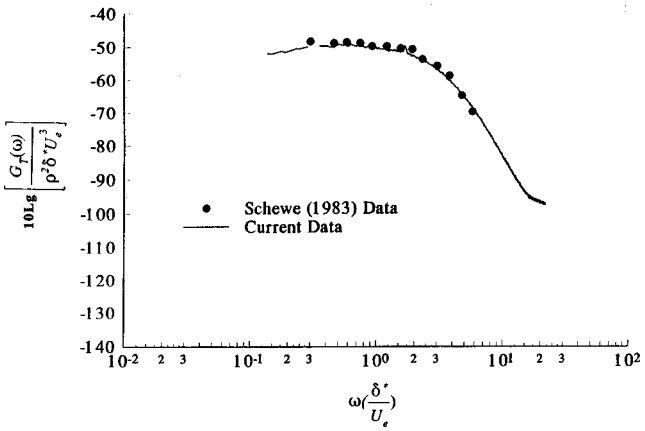


FIG. 18. Comparison of flat test plate boundary layer pressure spectrum with data from Schewe (Ref. 10), wall stationary, X=0.864 m, U_e =10.6 m/s, h_{ch}=5.40 cm.

spectrum. Flow structures in the boundary layer log region supposedly dominate the Fig. 14 spectrum from 528 to 1223 Hz. Higher frequencies arise from turbulent structures closer to the wall, in the boundary layer buffer region.

The plate flow modifications due to reducing h_{ch} from 2.6 δ to 1.4 δ occur exclusively in the outer portion of the external flow boundary layer, above $Z/\delta \approx 0.7$, where the velocity statistics are modified substantially. The spectra of Fig. 14 are almost identical except for a 1-dB difference over the lower range of frequencies measured. It appears that the flow modifications were insufficient to effect any significant change on the wall-pressure fluctuations.

The intermediate channel height stationary-wall results discussed above are consistent with the moving-wall data. The velocity data showed that wall motion tended to increase U_b by roughly 2%. For Z values where the flow was modified, the turbulence intensity levels fall roughly midway between the stationary wall results for $h_{ch}=2.6\,\delta$ and 1.4 δ. Therefore, the effect of wall motion on the pressure fluctuations measured for $h_{ch}/\delta=1.4$, at a given velocity, is expected to be negligible, as Fig. 15 supports.

Compared to the external boundary layer, the 0.7 δ channel-velocity statistics have been modified even more than those in the 1.4 δ channel, yet Fig. 14 shows that only at frequencies above 1200 Hz have the wall-pressure fluctuations been modified by more than 1 dB from the external-flow spectrum. These results demonstrate an insensitivity of the wall-pressure spectrum to the outer-flow structures that supposedly dominate the mid-range of frequencies.[16]

For the case of wall motion in the 0.7 δ channel, U_b is not only increased by 8%, but the flow characteristics are changed from approximately that of a fully developed channel flow to that of a moving-wall channel flow. The turbulence data and mathematical analysis in Brungart[17] indicate that the turbulence quantities for the stationary and moving-wall cases converge as the plate is approached and scale reasonably well on the local velocity. Figure 16 shows that the moving-wall spectrum is higher in level than the corresponding stationary-wall spectrum and is the same as the stationary wall-pressure spectrum measured at the value of the average velocity, U_b. The pressure fluctuations are essentially independent of the details of the outer flow.

The insensitivity of the currently discussed wall-pressure fluctuations to the specifics of the outer flow is further demonstrated in Fig. 19, which shows the nondimensionalized spectra, compared with a similar spectrum measured by Lauchle and Daniels[12] in fully developed pipe flow, and to an external turbulent boundary layer spectrum measured by Schewe.[10] The outer flows under which these three spectra were measured are distinctly different, yet the wall-pressure spectra are almost identical. Spectra measured at the same momentum-thickness Reynolds number have the same ratio of inner-to-outer flow time and amplitude scales. They can, therefore, be compared across the entire frequency range when nondimensionalized with either inner- or outer-flow variables. The momentum-thickness Reynolds number of the Fig. 19 data range from 1120 to 1800.

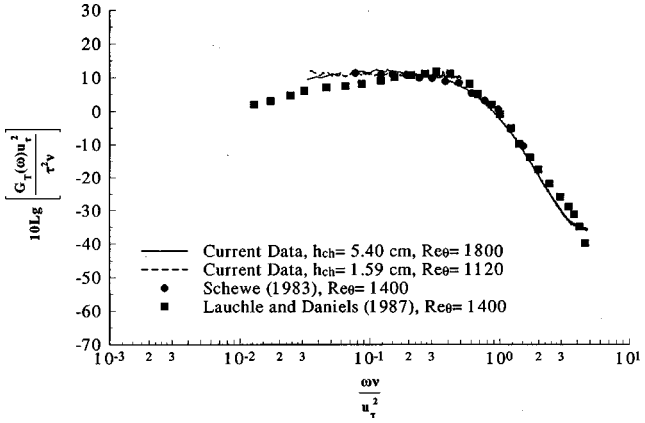


FIG. 19. Comparison of turbulent boundary layer wall-pressure fluctuation spectra normalized with inner-flow variables and measured in different facilities, or under different outer-flow conditions, but all at comparable values of Re_θ.

IV. SUMMARY

To better understand the outer-flow contribution to the pressure fluctuations occurring at the wall beneath a turbulent boundary layer, we have modified the outer flow of a flat plate boundary layer and compared the velocity and pressure fluctuation statistics with those of the unmodified outer-flow case. The modification was performed by setting one wall of a channel into streamwise motion. The height of the test plate above the moving (or stationary) wall was varied such that interacting and noninteracting turbulent boundary layers were included in the test matrix. The moving tunnel wall is effective in modifying the channel mean velocity, turbulence levels, Reynolds stresses, and, depending on the channel height, the wall-pressure fluctuations. This work has also demonstrated the relative insensitivity of the wall-pressure fluctuations to the details of the outer flow, even at those frequencies that are considered to be influenced by outer-flow turbulence structures.

ACKNOWLEDGMENTS

This work was conducted at the Applied Research Laboratory of the Pennsylvania State University through a grant from Ford Motor Co. This support, including that of Dr. Gary S. Strumolo, the project monitor at Ford, is gratefully acknowledged.

[1] R. H. Kraichnan, ''Pressure Fluctuations in Turbulent Flow Over a Flat Plate,'' J. Acoust. Soc. Am. **28**, 378–390 (1956).
[2] G. M. Lilley and T. H. Hodgson, ''On Surface Pressure Fluctuations in Turbulent Boundary Layers,'' AGARD Report No. 276, April 1960.
[3] R. L. Panton and J. H. Linebarger, ''Wall Pressure Spectra Calculations for Equilibrium Boundary Layers,'' J. Fluid Mech. **65**, 261–287 (1974).
[4] W. C. Meecham and M. T. Tavis, ''Theoretical Pressure Correlation Functions in Turbulent Boundary Layers,'' Phys. Fluids **23**, 1119–1131 (1980).
[5] D. M. Chase, ''Modeling the Wavevector-Frequency Spectrum of Turbulent Boundary Layer Wall Pressure,'' J. Sound Vib. **70**, No. 1, 29–67 (1980).
[6] D. M. Chase, ''The Character of the Turbulent Wall Pressure Spectrum at Subconvective Wavenumbers and a Suggested Comprehensive Model,'' J. Sound Vib. **112**, 125–147 (1987).
[7] W. W. Willmarth and C. E. Wooldridge, ''Measurements of the Fluctuat-

ing Pressure at the Wall Beneath a Thick Turbulent Boundary Layer,'' J. Fluid Mech. **14**, 187–210 (1962).
[8] M. K. Bull, ''Wall-Pressure Fluctuations Associated with Subsonic Turbulent Boundary Layer Flow,'' J. Fluid Mech. **28**, 719–754 (1967).
[9] W. K. Blake, ''Turbulent Boundary Layer Wall Pressure Fluctuations on Smooth and Rough Walls,'' J. Fluid Mech. **44**, 637–660 (1970).
[10] G. Schewe, ''On the Structure and Resolution of Wall-Pressure Fluctuations Associated with Turbulent Boundary-Layer Flow,'' J. Fluid Mech. **134**, 311–328 (1983).
[11] T. M. Farabee, ''An Experimental Investigation of Wall Pressure Fluctuations Beneath Non-Equilibrium Turbulent Flows,'' David W. Taylor Naval Ship Research and Development Center Report No. DTNSRDC-86/047 (1986).
[12] G. C. Lauchle and M. A. Daniels, ''Wall Pressure Fluctuations in Turbulent Pipe Flow,'' Phys. Fluids **30**, 3019–3024 (1987).
[13] W. L. Keith, D. A. Hurdis, and B. M. Abraham, ''A Comparison of Turbulent Boundary Layer Wall Pressure Spectra,'' J. Fluids Eng. **114**, 338–347 (Sept. 1992).
[14] W. L. Keith and J. C. Bennett, ''Low Frequency Measurements of the Wall Shear Stress and Wall Pressure in a Turbulent Boundary Layer,'' AIAA J. **29**, No. 4, 526–530 (1991).
[15] R. L. Panton, A. L. Goldman, R. L. Lowery, and M. M. Reischman, ''Low-Frequency Pressure Fluctuations in Axisymmetric Turbulent Boundary Layers,'' J. Fluid Mech. **97**, Part 2, 299–319 (1980).
[16] T. M. Farabee and M. J. Casarella, ''Structural Features of Wall Pressure Fluctuations Beneath Turbulent Boundary Layers,'' Phys. Fluids A **3**, No. 10, 2410–2420 (1991).
[17] T. A. Brungart, ''Boundary Condition Effects on Turbulent Boundary Layer Wall Pressure Fluctuations,'' Ph.D. thesis, Penn State University, 1997.
[18] J. H. Prout and R. C. Marboe, ''ARL Penn State Flow-Through Anechoic Chamber,'' ARL Penn State Technical Memorandum 89–65, 22 February 1990.
[19] R. C. Marboe, G. C. Lauchle, and W. A. Kargus IV, ''Quiet Wall Jet Facility for Basic Aero/Hydroacoustics Research,'' NCA-Vol. 10, *Hydroacoustics Facilities, Instrumentation and Experimental Techniques*, ASME, 1991.
[20] M. R. Davis, ''Design of Flat Plate Leading Edges to Avoid Flow Separation,'' AIAA J. **18**, No. 5, 598–600 (1980).
[21] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics*, 3rd ed. (Wiley, New York, 1982), p. 227.
[22] I. C. Lekakis, R. J. Adrian, and B. G. Jones, ''Measurement of Velocity Vectors with Orthogonal and Non-orthogonal Triple-Sensor Probes,'' Exp. Fluids **7**, 228–240 (1989).
[23] J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures*, 2nd ed. (Wiley, New York, 1986).
[24] L. R. Purtell, P. S. Klebanoff, and F. T. Buckley, ''Turbulent Boundary Layers at Low Reynolds Numbers,'' Phys. Fluids **24**, 802–811 (1981).
[25] D. E. Coles, ''The Law of the Wake in the Turbulent Boundary Layer,'' J. Fluid Mech. **1**, 191–226 (1956).
[26] P. S. Klebanoff, ''Characteristics of Turbulence in a Boundary Layer with Zero Pressure Gradient,'' NACA Technical Note 3178 (1954).
[27] R. L. Panton and J. M. Miller, ''Excitation of a Helmholtz Resonator by a Turbulent Boundary Layer,'' J. Acoust. Soc. Am. **58**, 800–806 (1975).
[28] H. Tennekes and J. L. Lumley, *A First Course in Turbulence* (MIT Press, Cambridge, MA, 1972).
[29] G. M. Corcos, ''Resolution of Pressure in Turbulence,'' J. Acoust. Soc. Am. **35**, 192–199 (1963).
[30] E. R. Lindgren, ''Experimental Study on Turbulent Pipe Flows of Distilled Water,'' Oklahoma State Univ., Civil Eng. Dept., Report 1 AD621071 (1965).

Effects of absorptivity due to fish on transmission loss in shallow water

Orest Diachok
Naval Research Laboratory, Washington, DC 20375

(Received 29 September 1997; revised 23 October 1998; accepted 29 October 1998)

Absorption losses at the resonance frequencies of sardines, 1.3 kHz (18 dB) at night, 1.7 kHz (15 dB) during the day and 2.7 kHz (35 dB) at dawn, were observed at a range of 12 km at a shallow water site in the Gulf of Lion in September of 1995. These observations were made during Modal Lion, a multidisciplinary experiment, which was designed to isolate absorptivity due to fish from other effects on long range propagation. Systematic changes in the resonance frequency of dispersed sardines were consistent with concurrent echo sounder observations of the vertical migration of sardines at twilight. Comparison of transmission loss measurements with a numerical sound propagation model that incorporates absorption layers in the water column permitted estimation of the average absorption coefficient, depth and thickness of absorption layers. Depths and thickness of layers estimated from sound propagation measurements were in good agreement with echo sounder data. Measured resonance frequencies, f_0, of dispersed fish were in good agreement with theoretical computations based on measured swim bladder dimensions. The measured resonance frequency of sardines in schools, which were at a depth of about 65 m, was about 0.6 times the resonance frequency of dispersed sardines. The observed frequency shift is consistent with an analytical equation of the fundamental resonance frequency of a ''cloud'' of bubbles, an average separation between fish in schools of about 0.8 fish lengths and 5×10^3 fish per school. These estimates are consistent with previously published estimates of these parameters. The resonance frequencies of other absorption lines are consistent with the hypothesis that absorptivity due to sardines generally manifests itself in pairs of absorption lines, which correspond to the resonance frequencies of an ensemble of individual sardines, and the resonance frequencies of an ensemble of schools. Matching of theoretical calculations with inferences of the absorption coefficient permitted estimation of the average number of dispersed sardines per unit volume, which also consistent with previously published estimates. The results presented here suggest the possibility of long term tomographic mapping of fish parameters over large areas using broadband transmission loss measurements. © *1999 Acoustical Society of America.* [S0001-4966(99)04703-7]

PACS numbers: 43.30.Es [SAC-B]

INTRODUCTION

Extinction of sound propagating through aggregations of pelagic fish has been investigated experimentally at high frequencies (far above swim bladder resonance) in a few controlled experiments summarized by Ye (1996). The objective was to quantify the effects of extinction on estimation of fish concentration using ship-mounted, downward-looking echo sounders. At 30–200 kHz high concentrations of fish at relatively shallow depths can effectively ''shadow'' fish at greater depths and the bottom. Weston (1967, 1970), and Ching and Weston (1971) reported measurements of sound extinction due to fish at resonance frequencies. The frequency range of their experiments was between 0.3 and 4.4 kHz. Differences between average nighttime and daytime signal levels were measured on a bottom-mounted hydrophone located tens of km from bottom-mounted sources in shallow water over a period of several years. Attenuation due to pelagic fish was hypothesized to be high at night, when fish are dispersed near the surface, and zero during the day, when fish are in schools near the bottom. On some days (hypothetically when fish concentrations were relatively high) the average day–night difference in signal levels was as high as 50 dB. Since the publication of these observations, there have been no reported experiments designed to quantitatively test and demonstrate the relationships between the bioacoustic parameters of fish and extinction coefficients at resonance frequencies, despite its potential importance for naval operations and biomass estimation.

The primary determinants of sound propagation in shallow water are the sound speed profile and the bottom critical angle. A source placed below the thermocline excites predominantly low order modes, which dominate signal levels recorded on hydrophones placed below the thermocline. Calculated signal levels due to low order modes above the thermocline are relatively low. Previously reported measurements of transmission loss below the thermocline are in accord with deterministic computations. Measured signal levels above the thermocline, however, are much higher than predicted by adiabatic mode computations. High energy levels in this ''shadow zone'' can be simulated with propagation models that permit scattering of low to high order modes by bottom roughness (Rouseff and Ewart, 1995) or internal waves propagating along the thermocline (Thiele, 1990; Schneider, 1990; Sellschopp, 1990).

In this paper a deterministic model will be employed to compute acoustic fields below the thermocline, to permit in-

ference of the effects of fish schools near the bottom on low order modes. A heuristic model, in which replica fields are derived from measured signal levels on hydrophones above the thermocline during the day, when no fish are present above the thermocline, will be employed to permit inference of the effect of dispersed fish near the surface at night on high order modes. The latter procedure is analogous to the procedure adopted by Ching and Weston in their analysis of their data. In this paper measurements which were made on a hydrophone array, which spanned most of the water column, are analyzed to estimate the effects of both dispersed fish near the surface at night and schools of fish which reside near the bottom during both night and day.

This paper is organized into 14 sections. The objective of Sec. I is to summarize the bioacoustic properties of pelagic fish with swim bladders. Section II provides a description of a multidisciplinary experiment designed to isolate the effects of fish-related absorptivity on propagation loss at resonance frequencies. The environmental scenario where this experiment took place is described in Sec. III. The resultant measurements are described in Sec. IV. Section V provides a description of a deterministic transmission loss model that includes source directionality. The effects of absorbing layers on transmission loss are illustrated in Sec. VI. The inferred frequency dependence of absorptivity due to layers of fish schools near the bottom, where deterministically predictable, low order modes dominate, are described in Sec. VII. Section VIII provides a description of the properties of forward scattered energy evident above the thermocline. The inferred frequency dependence of the absorptivity due to layers of dispersed sardines above the thermocline, where forward scattered energy dominates, is presented in Sec. IX. The inferred frequency dependence of the absorption coefficient of dispersed fish at dawn, just prior to school formation, and during daytime are considered in Sec. X. Section XI provides a discussion of the downward shift in resonance frequencies of sardines in schools. The magnitude of Q_0 (defined as the ratio of the resonance frequency and spectral width of individual sardines), a parameter required for calculation of number density, is discussed in Sec. XII. Section XIII provides an estimate of the average number of sardines per m^3 along the measurement track. The paper concludes with Sec. XIV, a summary of the most significant results, conclusions and suggestions for future research.

I. BIOACOUSTIC PARAMETERS

Two types of pelagic fish dominate the Mediterranean Sea: sardines and anchovies (Anonymous, 1981). Their lengths are approximately 16 and 13 cm, respectively. Both fish are physotomes. Such fish have a pneumatic duct between the gut and the swim bladder, permitting passage of air swallowed at or near the surface. Physotomes do not have a gland that secretes gas into the swim bladder. The volume of the swim bladder of physotomes varies inversely with pressure. Physoclists are a class of fish that has a gland that secretes gas into the swim bladder. The volume of the swim bladder of physoclists is independent of depth (Blaxter and Batty, 1990). Measurements of the depth dependence of the resonance frequency of target strength indicate that the majority of fish in the ocean fall into one of these two categories (Hersey *et al.*, 1962).

A. Resonance frequency of individual fish

Andreeva (1964) proposed that the resonance frequency, f_0, of swim bladders of individual fish may be described with the equation:

$$f_0=(1/2\pi r)(3\gamma+4\mu_1)^{1/2}p^{1/2}/\rho_0^{1/2}, \tag{1}$$

where r is the radius of the bubble, p is the pressure, ρ_0 is the density of sea water, γ is the ratio of specific heats of the gas in the swim bladder, and μ_1 is the real part of the shear modulus of fish flesh. A recent review of the accuracy of input parameters into Eq. (1) is provided by Ye and Farmer (1994). According to Love (1978), measurements of f_0 of physotomes are approximately equal to theoretically computed values of f_0 with μ_1 equal to zero. The resultant equation has the same form as the equation for f_0 of an ideal bubble in water (Minnaert, 1933). Love also noted that measurements of f_0 of physoclists can be up to several times larger than computations based on Eq. (1) (with $\mu_1=0$), and formulated an alternate model to account for these observations. This paper is restricted to analysis of measurements made on physotomes. Substituting $r=r_0(p_0/p)^{1/3}$ and $p=p_0(1+0.1d)$, permits Eq. (1) to be rewritten as

$$f_0=(1/2\pi r_0)(3\gamma/\rho_0)^{1/2}p_0^{1/2}(1+0.1d)^{5/6}, \tag{2}$$

where r_0 is the radius of the swim bladder at the surface, p_0 is the pressure at the surface, and d is the depth in meters. This equation neglects the effects of temperature on r. The resultant error is less than 1%. The swim bladder is assumed to be filled with air, consequently, $\gamma=1.40$, and the effective density is assumed to be the density of sea water, $\rho_0=1.026\ \mathrm{gm/cm^3}$.

The swim bladders of fish are elongated. The geometrical shapes of swim bladders may be approximated with prolate spheroids or cylinders with end caps (Feuillade and Werby, 1994; Clay, 1991). The effective radius, r_0, of swim bladders in Eq. (1) is defined as the radius of a sphere which has the same volume as the swim bladder (Holliday, 1972).

The size of the swim bladder increases as fish grow. The relationships between r_0 and the length, L, of anchovies (Holliday, 1978) and pilchard sardines (Scalabrin, personal communication) are shown in Fig. 1 (top). Calculations are provided for both geometrical models of swim bladders (the top set of data points correspond to cylinders). The empirical relationship between these parameters for anchovies is based on 21 samples taken off the coast of California (Holliday, 1978). Scalabrin measured L and r_0 on 15 pilchard sardines taken from the Bay of Biscay. These data suggest that r_0 increases rapidly with L when L is small, and slowly with L when L is large. Measurements on 8–27 cm long coalfish (Lovik and Hovem, 1979), and 3–25 cm long herring (Blaxter and Batty, 1990) exhibit a similar behavior.

The resonance frequency of an elongated bubble is significantly greater than the resonance frequency of a spherical bubble with the same volume (Feuillade and Werby, 1994; Clay, 1991). The scaled increase in resonance frequency, x,

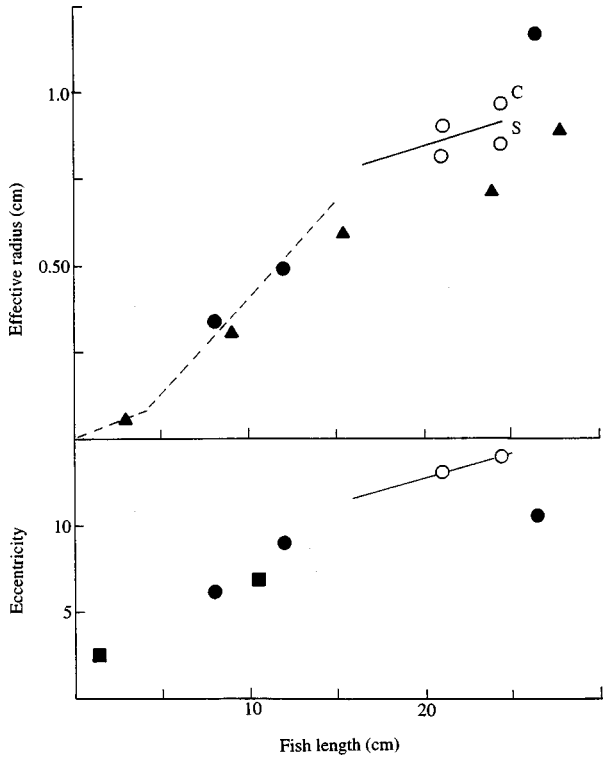


FIG. 1. Relationship between the effective radius, r_0 (top), and eccentricity, ϵ (bottom), of swim bladders at the surface and fish length, L, for pilchard sardines (○) due to Scalabrin, coalfish (●) due to Lovik and Hovem (1979), and herring (▲) due to Blaxter and Batty (1990), and r_0 (– –) and ϵ (■) of anchovies due to Holliday (1978). Calculated values of r_0 of sardines are based on cylindrical (*C*) and spheroidal (*S*) models of the swim bladder. Also shown: best fit through modal values derived from Scalabrin's measurements (——).

is a function of the eccentricity (the ratio of the major to minor axes), ϵ, of swim bladders. The modal values of ϵ of 21 and 24.5 cm long sardines (which are shown in Fig. 1, bottom), are equal to 13 and 14, respectively (Scalabrin, personal communication). The extrapolated value of ϵ of 16.5 cm long sardines is 12. For comparison, the average values of ϵ of <3 cm long and 6–15 cm long anchovies are 2.6 (at night) and 7 cm, respectively (Holliday, 1978). These data suggest that ϵ increases systematically with L. Measurements of ϵ of 8–27 cm long coalfish (Lovik and Hovem, 1979), which are also shown in Fig. 1 (bottom), exhibit a similar behavior. Measurements of ϵ of 1.5 cm long anchovies indicate that ϵ increases with depth (Holliday, 1978). It is not known if ϵ of adult sardines is depth dependent. According to calculations based on the prolate spheroid and cylindrical models, x is 1.28 and 1.34 for 16.5 cm long sardines respectively. To account for this effect, Eq. (2) should be modified to

$$f_0=(x/2\pi r_0)(3\gamma/\rho)^{1/2}p_0^{1/2}(1+0.1d)^{5/6}, \tag{3}$$

and further simplified to

$$f_0=322x(1+0.1d)^{5/6}/r_0, \tag{4}$$

where r_0 is in cm, and f_0 is in Hz.

Controlled measurements of f_0 of anchovies (Baltzer and Pickwell, 1970), char (Sundness and Sand, 1975), sprat and herring (Lovik and Hovem, 1979) were 3%, 1%, 7% and 22% respectively, higher than calculated values of f_0 based on Eq. (4). The latter discrepancy may be due to physiological effects associated with high oil content (Blaxter and Batty, 1990), or to swim bladder compression due to an enlarged gonad or a full stomach. Ona's (1990) laboratory measurements of the volumes of cod swim bladders suggest that enlarged gonads can reduce r_0 by a factor of 0.85 (compared to r_0 of a ''normal'' fish) and significantly increase the magnitude of ϵ. His measurements also indicate that a nearly (3/4) full stomach can reduce r_0 by a factor of 0.79 (compared to r_0 of a fish with an empty stomach) and significantly increase the magnitude of ϵ. Theoretically decreases in r_0 and increases in ϵ lead to higher values of f_0. The first phenomenon may be expected to have the most pronounced effect on f_0 during a few months before and during the peak spawning month (January in the Gulf of Lion). Measurements, which will be presented in this paper, were made in early September, when spawning related effects are probably small. The effect of the second phenomenon is probably small, except possibly during the spring bloom, which usually coincides with the peak spawning month (Blaxter and Hunter, 1982). According to Loukashkin (1970), the majority (~70%) of stomachs of anchovies, which were randomly sampled at sea, were essentially empty. Only about 10% were more than 75% filled. Nonacs *et al.*'s (1998) model suggests that this type of distribution is also representative of sardines.

B. Extinction cross section and absorption coefficient

At the resonance frequency the extinction cross section, σ_E, is related to the scattering cross section, σ_S (Clay and Medwin, 1977):

$$\sigma_E=\sigma_S\lambda_0/2Q_0r=2rQ_0\lambda_0, \tag{5}$$

where Q_0 equals $f_0/\Delta f_0$ and Δf_0 is the spectral width (measured at 3 dB below peak levels), and λ_0 is the wavelength corresponding to f_0.

The absorption coefficient, α_0, due to a large number of identical fish at f_0 (Clay and Medwin, 1977) may be expressed as

$$\alpha_0=1/2n\sigma_E, \tag{6}$$

where n is the number of fish per unit volume.

C. Geometrical properties of fish schools

The objective of this section is to review the geometrical properties of schools, which are input parameters for theoretical models of the acoustic properties of fish schools. Pelagic fish in the ocean occur in one of three idealized modes: dispersed, in schools, and in shoals. Following Pitcher and Parrish (1993), schools are defined as structures in which fish are closely spaced, polarized and synchronized; schools generally consist of fish of the same species and similar size (Pitcher *et al.*, 1985). Fish form schools in response to predators. Shoals are defined as structures in which fish are widely spaced, randomly oriented, and loosely organized. Shoals are usually found near the surface at night. Fish form shoals

when they are foraging for prey. Nonacs *et al.* (1998) provide the following hypothesis of the transition process between these structures: ''A fish school upon entering a zooplankton patch expands to fill the entire volume (of the patch)...and after feeding, contracts to its original size and exits the patch.'' According to Smith *et al.* (1989), the modal diameter and thickness of plankton patches are about 500 and 20 m, respectively. Nonacs *et al.*'s (1998) postulate implies that schools and shoals coexist at the same depth.

The internal structure of schools of herring has been mapped with a high resolution sonar by Misund (1991). His measurements reveal that schools consist of an inner core region (the ''nucleus'' of the school), in which s is usually equal to approximately L, and a peripheral region (the ''fuzz'' of the school), in which s is significantly larger than L. The ''nucleus'' constitutes about half of the volume of schools.

This result is qualitatively consistent with a large set of three-dimensional photographic images of anchovies in shoals at night, which were recorded by Aoki and Ingaki (1989). Their results indicated that the modal value of s was relatively small ($2L$ and $3L$ at depths of 40 and 20 m, respectively). Their measurements also indicated that the majority of separations are much larger than L. Analysis of photographic images of anchovy schools in the ocean during daytime (Graves, 1977), revealed that the modal value of s was about $0.8L$. His measurements also revealed that the majority of separations were significantly larger than L.

Freon *et al.*'s (1996) measurements of the number density of sardine schools, which were made with a downward-looking echo sounder, revealed that the average values of s of schools were $3L$ at 12 m at night, and $2L$ at 20 m during daytime. Modal values of s were not reported, but were probably smaller. Evidently, s depends on depth, light intensity and probably on the kind and concentration of predators and prey (Blaxter and Hunter, 1982). The average distance between ''nearest neighbors'' in pilchard sardine schools, which was derived from three-dimensional analysis of photographs made under controlled conditions in a large pool during daytime, was about $0.6L$ (Cullen *et al.*, 1963). For sardines encountered in this experiment, a separation of L corresponds to 16.5 cm. The resonance frequency of 16.5 cm long sardines at 65 m, the depth of sardine schools during this experiment during daytime, is approximately 2.8 kHz. The corresponding wavelength is 0.5 m; consequently s/λ_0 equals approximately 0.3 in schools in which s equals L, and approximately 3.0 in regions where s equals $10L$.

It will be shown in Sec. I F that the resonance frequency of the ''nucleus'' of schools is smaller than the resonance frequency of individual fish when $s/\lambda_0 < 1$ (where λ_0 is the wavelength at f_0). Misund's results suggest that absorption coefficients associated with a layer of ''schools'' should exhibit two absorption lines, which are due to individuals and the ''nucleus'' of schools. For individuals in the ''fuzz'' or periphery of schools, $s/\lambda_0 > 1$. For ''nuclei'' of schools, $s/\lambda_0 < 1$.

According to Scalabrin *et al.* (1996), and Scalabrin (1997) the modal horizontal and vertical dimensions of schools of adult sardines during daytime are 6 and 2 m, respectively. Average dimensions of schools are larger. These measurements were made with a downward-looking echo sounder. There was no attempt to discriminate between ''schools'' and ''nuclei of schools.'' Consequently, the dimensions of school ''nuclei'' may be smaller.

Schools can assume a variety of shapes (Blaxter and Hunter, 1982; Radakov, 1973), and to a first approximation, may be modeled as oblate spheroids (Misund, 1995). The effective radius, R, of an oblate spheroid with a diameter of 6 m and a thickness of 2 m is 2.08 m. The number of fish per school, N, may be calculated with the equation

$$N = 4/3(\pi R^3 n), \tag{7}$$

where n is the number density, which is equal to $1/s^3$. If the average separation between fish in schools is 0.165 m, then N equals approximately 0.8×10^3 fish per school.

D. Resonance frequency of fish in schools

The objective of this section is to summarize the results of Feuillade *et al.*'s (1996) and Nero's (1996) calculations of the dependence of the resonance frequency of schools and the scattering cross section per fish on the spacing to wavelength ratio, and the total number of fish per school. When the average separation between fish in a school (s) is comparable to $\lambda_0/2$, $f_0'/f_0 < 1$ (where f_0' is the resonance frequency of a school) and $\sigma_S'/\sigma_S < 1$ (where σ_S' is the scattering cross section of one fish within a school) (Feuillade, 1995). Scattering from an ensemble of bubbles may be theoretically described in terms of the normal modes of oscillation of the ensemble. Modes may be ''symmetric,'' where bubbles oscillate in phase with each other, or ''antisymmetric,'' where some or all of the bubbles oscillate in antiphase with each other. Symmetric modes typically show downward shifts in frequency and reduced scattering cross sections.

Using this formalism, Feuillade *et al.* (1996), and Nero (1996, personal communication) calculated the scattering cross sections versus frequency of ''virtual'' three-dimensional schools containing up to 300 fish. Figure 2 provides an example of Nero's unpublished calculations for 100 fish in a spherical school with s/λ_0 equal to 0.14. For this value of s/λ_0, f_0'/f_0 equals 0.8 and σ_S'/σ_S equals 0.3. Figure 3(a) shows values of σ_S'/σ_S, which were inferred from their computations for schools containing 10, 13, 100 and 123 fish. These computations are representative of simulated schools in which the standard deviation of swim bladder radii, Δ_r, varies between zero and 10%. Nero's (1996) calculations suggest that the magnitudes of σ_S'/σ_S and f_0'/f_0 are nearly independent of Δ_r for Δ_r up to about 20%. Results shown in Fig. 3(a) indicate that the magnitude of σ_S'/σ_S is minimum when s/λ_0 equals approximately 0.18. For comparison, the value of s/λ_0 at which σ_0'/σ_0 is minimum for N equal to 2 and 3, is only slightly different, viz., 0.12 (Feuillade, 1995). The large variance in the magnitudes of σ_S'/σ_S for fixed values of s/λ_0 indicates that many additional simulations are needed to precisely define the average dependence of σ_S'/σ_S on s/λ_0. Nevertheless, these computations suggests that the dependence of σ_S'/σ_S on N is weak, and that

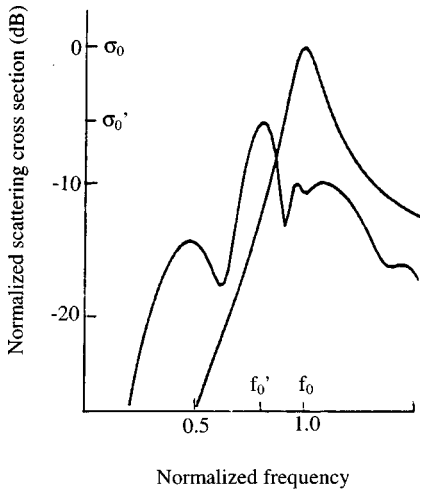


FIG. 2. Normalized scattering cross sections versus frequency for individual fish and for a spherical school in which the average separation, *s*, between 100 fish is one fish length, due to Nero. This figure illustrates the downward shift in resonance frequency, and the reduction of scattering cross section of fish in school. f_0 and f_0' are the resonance frequencies of individual fish and fish in school respectively. σ_0 at f_0 and σ_0' at f_0' are the scattering cross sections of dispersed fish (where $s>\lambda_0$) and fish in school (where $s<\lambda_0$).

they are representative of other (more realistic) values of *N*. These computations, together with Eq. (5), provide a means for estimating the magnitude of σ_E'/σ_E.

A summary of the magnitudes of f_0'/f_0, which were derived from Feuillade *et al.*'s computations for schools which contain 123 fish, and Nero's computations for schools which contain 100 fish, are plotted in Fig. 3(b) as a function of s/λ_0. These calculations indicate that f_0'/f_0 increases with s/λ_0, and approaches unity as s/λ_0 approaches 0.2. These calculations indicate that f_0'/f_0 increases approximately linearly with $N^{-1/3}$. This result is qualitatively consistent with analytical computations (Feuillade, 1995) for two and three

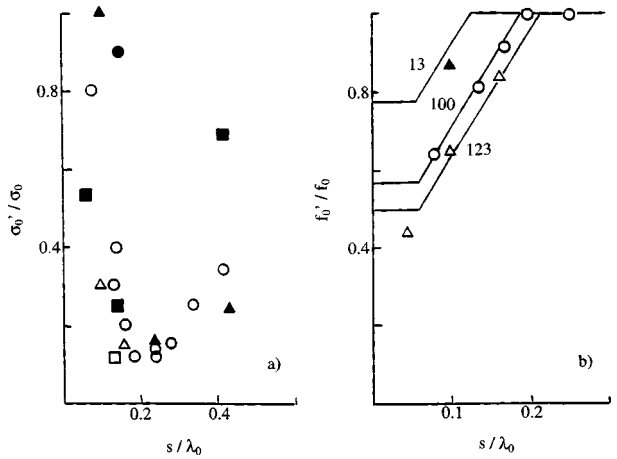


FIG. 3. (a) The magnitudes of σ_0'/σ_0 vs s/λ_0 (for various values of β), derived from simulations for *N* equal to 13 (▲) and 123 (△) fish per school, and ϵ equal to 1 (Feuillade *et al.*, 1996), for *N* equal to 100 (○) and ϵ equal to 1 (Nero, personal communication), and for *N* equal to 10 (●), 100 (■) and 300 (□), and ϵ equal to 5 (Nero, 1996); and (b) the magnitudes of f_0/f_0' vs s/λ_0 for *N* equal to 13 and 123, ϵ equal to 1 and β equal to 9.9×10^{-4}; and *N* equal to 100, ϵ equal to 1 and β equal to 5.2×10^{-4}. Horizontal lines are based on theoretical computations of f_0/f_0' of fundamental mode of vibration of spherical schools, when $s/\lambda_0\leq0.06$ [Eq. (10)]; diagonal lines are based on Eq. (12).

bubbles, which indicate that f_0'/f_0 decreases as the number of bubbles increases. The lines, which are shown in this figure, are based on equations that describe the resonance frequency of spherical ''bubble clouds,'' which will be discussed in the next section.

E. An analytical model of the resonance frequency of schools

The objectives of this section are to (1) describe a theoretical equation of the resonance frequency of bubble clouds, based on effective medium theory, and (2) formulate a correction to this equation to permit estimation of f_0'/f_0 for relatively large values of $s/\lambda_0(\sim0.3)$. Effective medium theory is applicable when (1) $s\ll R$, (2) $s\ll\lambda_0$, and (3) the effects of coupling between bubbles are negligible. An approximate theoretical equation, which is due to d'Agostino and Brennan (1988), will be employed to account for Nero's computations in the limit of small values of s/λ_0 and small values of *N*. This equation describes the normal modes of a spherical bubble cloud when $s/\lambda_0<0.06$. Their equation will then be modified to be consistent with his computations at higher values of s/λ_0 and small values of *N*. Finally, the modified equation will be exercised at relatively large values of $s/\lambda_0(\sim0.3)$ and large values of $N(\sim10^4)$ for comparison with measurements. According to d'Agostino and Brennan, the normal modes of a spherical bubble cloud may be expressed as:

$$f_0'/f_0\approx[1+(3\beta R^2)/(m-1/2)^2\pi^2r^2]^{-1/2}, \tag{8}$$

where *R* is the radius of the school, *m* is the mode number, and β is the void fraction, which is defined as:

$$\beta=(4/3)\pi r^3/s^3. \tag{9}$$

R, *s* and the number of fish per school, *N*, are related through Eq. (7). Based on Eqs. (7) and (9), Eq. (8) may be expressed as:

$$f_0'/f_0\approx[1+(3N^{2/3}\beta^{1/3})/(m-1/2)^2\pi^2]^{-1/2}. \tag{10}$$

For values of $[(3N^{2/3}\beta^{1/3})/(m-1/2)^2\pi^2)]$, which are large compared to 1, Eq. (10) reduces to Carey and Roy's (1993) and Lu *et al.*'s (1990) first order approximation for the fundamental mode of a spherical bubble cloud,

$$f_0'/f_0\approx N^{-1/3}\beta^{-1/6}. \tag{11}$$

Measurements performed on experimentally generated bubble clouds have confirmed the validity of Eq. (11) [and by inference, Eq. (10)] for s/λ_0 equal to 0.06 (Schindall, 1995).

The horizontal extent of schools is generally larger than their vertical extent. Consequently, the resonance frequency of schools will be larger than predicted by Eq. (11). To account for this effect, Eq. (10) should be modified to include a correction for the eccentricity of the assumed oblate spheroidal shape, x_S (Weston, 1967).

The horizontal lines shown in Fig. 3(b) are for *N* equal to 100 and β equal to 5.2×10^{-4}, and for *N* equal to 123 and β equal to 9.9×10^{-4}. For both computations, x_S equals 1. These values of β were derived from Feuillade *et al.*'s and Nero's assumed values of *r* and *s*. The inferred value of

f_0'/f_0 [shown in Fig. 3(b)] for s/λ_0 equal to 0.05, is equal to 0.44. This value is in good agreement (within 12%) with the computation based on Eq. (10), viz., 0.50.

For higher values of s/λ_0, the departure of Feuillade *et al.*'s and Nero's computations from Eq. (10) is much larger; presumably because the requirement, that s/λ_0 is small, is not strictly satisfied in the simulated schools. The difference between the simulations and Eq. (10) increases approximately linearly with s/λ_0, when s/λ_0 is larger than 0.06. The magnitudes of f_0'/f_0 for values of s/λ_0, which are larger than 0.06, are consistent with the following equation:

$$f_0'/f_0 \approx x_S[1+(12N^{2/3}\beta^{1/3}/\pi^2)]^{-1/2}[1+5.8(s/\lambda_0-0.06)]. \tag{12}$$

This equation reduces to Eq. (10) when s/λ_0 equals 0.06, and x_S equals 1, and encapsulates Feuillade *et al.*'s and Nero's computations for s/λ_0 greater than 0.06. Figure 3(b) also shows the results of Feuillade *et al.*'s computation of f_0'/f_0 for N equal to 13 and β equal to 9.9×10^{-4}, and analytical computations based on Eq. (12). The good agreement between these computations suggests that Eq. (12) provides a good approximation of the dependence of f_0'/f_0 on N, and may be employed for estimating the magnitude of f_0'/f_0 for large values of N.

In this experiment the magnitudes of λ_0 and r at a depth of 65 m are equal to approximately 0.52 m and 0.40 cm, respectively. If s is assumed to equal L (16.5 cm), then β equals 6.0×10^{-5}. If x_S is assumed to equal 1.12, then f_0'/f_0 [based on Eq. (12)] equals 0.51 for N equal to 1.6×10^4. Experimental estimates of f_0'/f_0, which were derived from Modal Lion data, will be discussed in Sec. XIII.

F. Diel variability

During the day pelagic fish on continental shelves generally occupy layers near the bottom. According to Azzali *et al.* (1985), the majority (~60%) form schools; the remainder are either dispersed (~30%) or in large scale nonhomogeneous ''accumulations,'' which consist of fish which are dispersed, in school, and in shoals (~10%). Since ''accumulations'' include schools, the percentage of fish in schools during daytime is generally between 60% and 70%. These figures represent year long averages; distributions between these modes on particular days may be substantially different and probably depend on the kind and concentrations of predators and prey and environmental conditions. At night most of these fish migrate upward and form shoals (Freon, 1996). In this mode fish occupy depths at which there is sufficient starlight, moonlight or bioluminescence to maintain visual contact, permit formation of schools and descend to greater depths when threatened (Smith, personal communication). According to Azzali *et al.* (1985), about 5% form schools, about 20% are dispersed, and about 75% occur in ''accumulations'' at night.

Evidently bioluminescence from plankton can provide sufficient light to permit schooling at or near the thermocline at night (Whitney, 1969). During the summer schools of sardines in the Adriatic Sea at night have been observed to occupy a layer which is about 20 m below a strong (~1 °C/m) thermocline (Zupanovich, 1967). Vertical migration between daytime and nighttime depths occurs over a period of about one hour before sunrise and after sunset (Azzali *et al.*, 1985; Freon *et al.*, 1996; Woodhead, 1966).

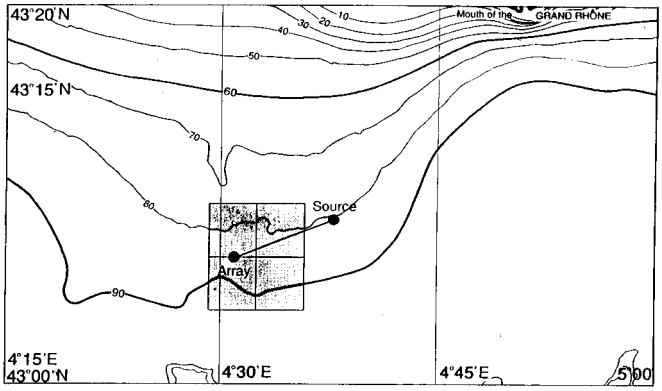


FIG. 4. Locations of source, receiving array, IFREMER survey site (shaded area) and depth contours in meters.

G. Concentrations

The average concentrations of sardines and anchovies have been thoroughly measured over the past 20 years in the Adriatic Sea (Azzali, 1995). Concentrations of these fish in this region are, on average, comparable to the Gulf of Lion, and may therefore be used to investigate the scales of temporal and spatial variability. Spatially averaged concentrations of sardines in the Adriatic Sea have varied between about 2 and 25 tons/km^2 over the past 20 years. Highly concentrated aggregations, or ''colonies'' of fish in this region are nominally 15 km in diameter, occupy about 15% of the area, and persist for several days. Concentrations within such colonies are significantly greater than spatially averaged concentrations. In particular, n/n_A, where n_A is the spatially averaged number density, and n is the number density in colonies, equals 4 in 15 km patches which occur in ~15% of the total area. For comparison, n/n_A equals 5 in ~12% of the total area, and 7 in ~5% of the total area. The biomass in colonies has varied between 8 and 100 tons/km^2 (assuming that n/n_A equals 4). Colonies, which have similar dimensions, have been observed in the northeast Pacific (Smith, 1978); in the Bay of Biscay (Diner, personal communication) and in the Gulf of Lion (Liorzou, personal communication). The temporal and spatial scales of colonies are hypothetically related to high concentrations of plankton associated with oceanographic features such as upwelling sites, fronts and eddies (Mullin, 1993).

II. EXPERIMENTAL DESIGN

Modal Lion was conducted in the Gulf of Lion between 1 and 10 September, 1995 at the site shown in Fig. 4. Fish concentrations at this site were relatively high based on a survey conducted from the R.V. L'EUROPE by Liorzou (personal communication). The survey included high frequency echo sounder measurements, and direct sampling of bioacoustic parameters of fish in the region. The main components of Modal Lion, which are illustrated in Fig. 5, included a mechanically steerable 40 kHz, 0.8 m diameter parametric

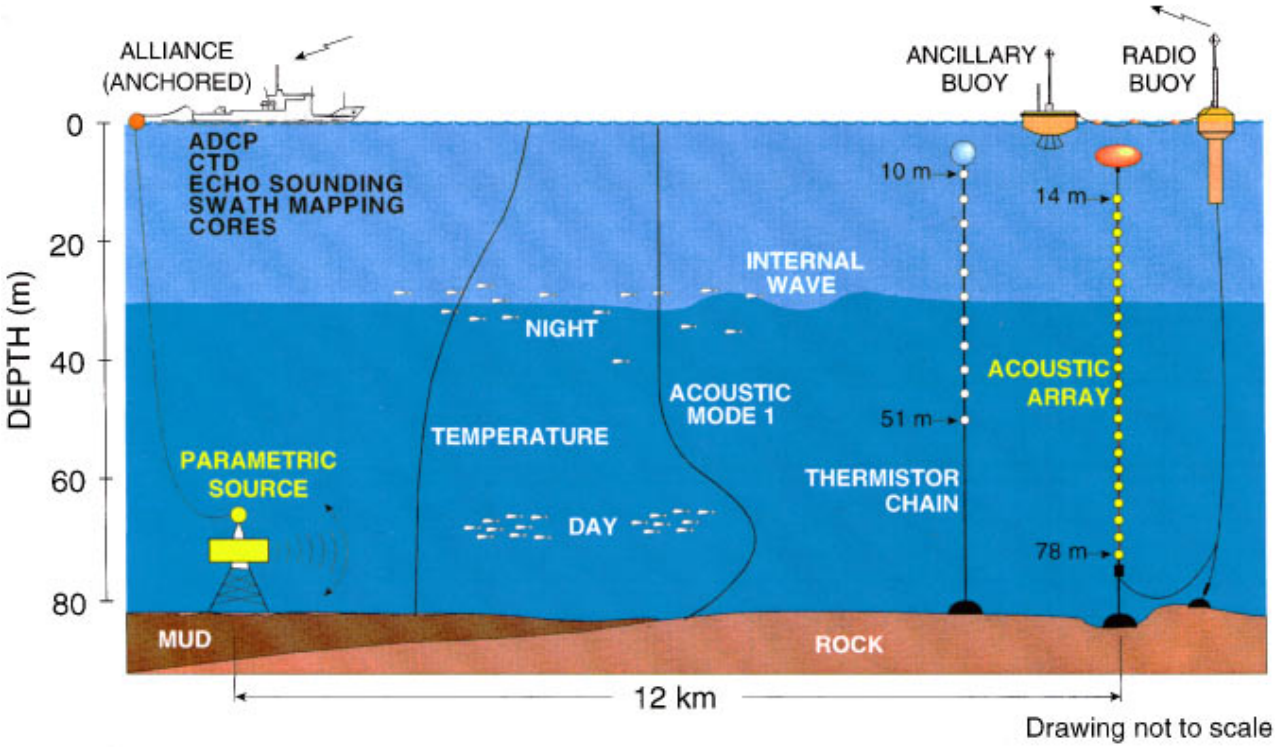


FIG. 5. Experimental geometry.

source (Moffet and Mellen, 1977), deployed 5 m above the bottom, and a vertical array that spanned most of the water column. Power to the parametric source was provided from the R.V. ALLIANCE, which was four point moored. Batteries and a diesel generator provided power to the vertical array; both were housed in an ancillary buoy. The vertical array was deployed in a slight (~4 m) depression in the bottom to ensure that the array, together with its buoy fit into the water column, and to permit sampling of the acoustic field over the majority of the water column. CW signals were transmitted every 15 seconds for periods of 2 seconds at frequencies, which were separated by approximately 10% increments with a repetition rate of 5 minutes for periods lasting 7 hours centered around sunrise and sunset. The transmitted frequencies were: 0.6, 0.7, 0.8, 0.9, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 2.2, 2.4, 2.7, 3.0, 3.4, 3.8, 4.2, 4.6, and 5.0 kHz. Ambient noise levels were measured at each frequency prior to signal measurements. Received energy levels were successfully recorded on 5, 6 and 10 September. The analysis presented in this paper will be limited to data recorded on 5 and 6 September.

To permit theoretical modeling of the results, conductivity and temperature versus depth profiles (CTD's) were measured near the source from the Alliance once per hour, and temperature was measured near the receiving array with a thermistor string every 5 minutes. Current speeds versus depth were measured from the ALLIANCE with a hull-mounted acoustic doppler current profiler (ADCP) every 5 minutes; wind speed and direction were recorded every 5 mins; and radar was employed to track fishing and other vessels.

The source levels of the parametric source were measured in January 1996 at a shallow (89 m) water site near La Spezia, Italy with the same equipment, geometrical configuration, signal transmission and signal processing procedures employed during Modal Lion. During these measurements the environment was essentially isothermal. The parametric source was tilted up a few degrees to minimize interaction of the source generation region with the bottom. Signals were received at ranges of 900 and 1800 m. Source level measurements made at these ranges were in agreement within ± 1 dB, confirming that these ranges were beyond the source generation region. The measured source level dependence on input voltage was in accordance with theory. Source levels were determined by matching received levels versus depth with normal mode computations that included source directionality. Coherent summation of modes provided good matches between theoretical and measured interference patterns at frequencies between 0.7 and 1.8 kHz. At higher frequencies incoherent summation of modes provided a better match between theoretical and measured energy levels at all depths. The loss of predictability of intermodal phase at the higher frequencies may have been due, in part, to differences between the actual and assumed locations of the hydrophones in the array (hydrophone locations were not measured), and in part, to ''microstructure'' in sound speed profiles. The uncertainty in the inferred source levels is approximately ± 1 dB at frequencies below 1.8 kHz, and approximately ± 2 dB at higher frequencies.

III. THE ENVIRONMENT

A. Bioacoustics

Echo sounder measurements of the distributions of fish depth were made with the R.V. ALLIANCE in moving and moored modes, prior to, during, and after the propagation

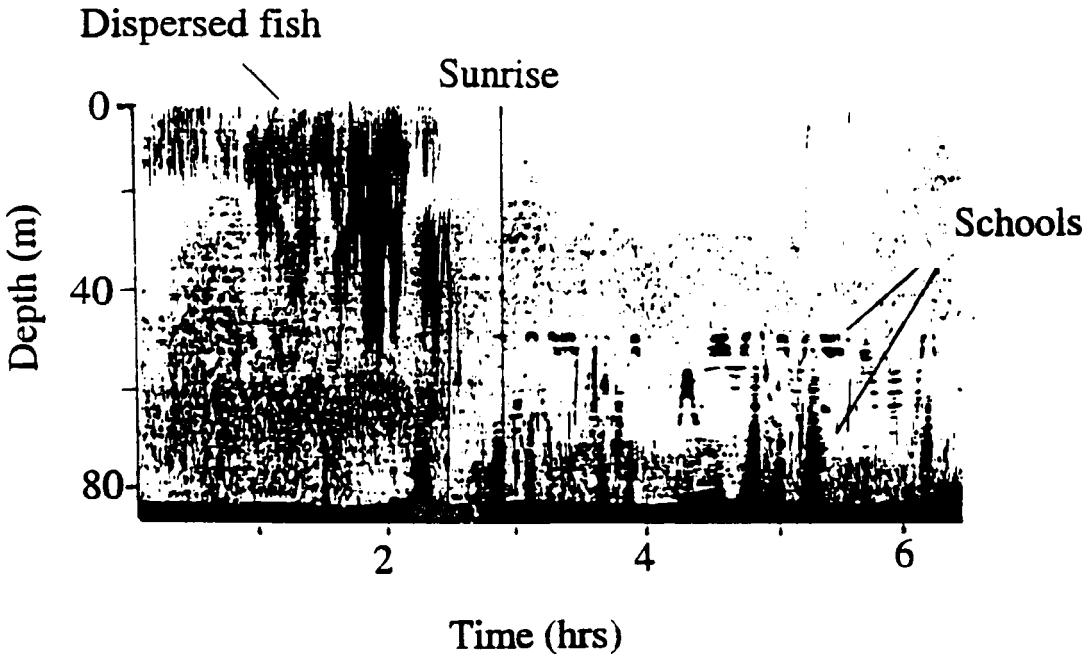


FIG. 6. Echo sounder record of the temporal evolution of the depth dependence of fish (predominantly sardines) and their transition from (1) predominantly dispersed mode at night, to (2) predominantly school mode during day. The vertical line marks sunrise.

experiments. An example of data recorded during 6 September, when the ship was in moored mode, is shown in Fig. 6. Schools of fish are clearly evident near the bottom during the day. Records at night exhibit the presence of a smaller number of schools near the bottom, and diffuse layers near the surface, which consist of dispersed sardines and anchovies.

Previously reported measurements indicate that dispersed sardines usually occur at depths which are between 15 and 35 m (Culley, 1971; Barange *et al.*, 1996), and that dispersed anchovies usually occur at depths which are ~9 m (Barange *et al.*, 1996) and <10 m (Hunter, personal communication). These estimates may be biased by avoidance of ship radiated noise and echo sounder transmissions (Soria, 1996; Misund and Aglen, 1992; Popper, 1997).

Prior to and after the propagation measurements, the R.V. ALLIANCE surveyed the track between the source and receiver. The observed mean depth of dispersed fish was 35 m. When the ship was stationary the mean depth of dispersed fish was approximately 25 m; hence, the actual depth was probably about 25 m. The layer attributed to anchovies (~9 m) was not evident in the echo sounder data. The modal depths of schools, which were derived from measurements made when the ship was stationary, were approximately 65 m deep during the day and at night. These measurements are representative of only the largest schools. Due to limitations imposed by the dynamic range of the recorder, these measurements are not suitable for inference of the total number of schools, but may be used to estimate the average depth and thickness of the layers.

Normalized distributions of the number of fish per km per 10 m depth interval versus depth, which were derived from measurements made when the ship was moving on 4 and 7 September during day and night, are illustrated in Fig. 7. The integrated number of schools per km during day and night were approximately equal. These data show that fish were concentrated in schools in a layer that was approximately 65 m deep and 15 m thick during the day, in agreement with measurements made in stationary mode; and in a layer that was about 65 m deep and 20 m thick during the

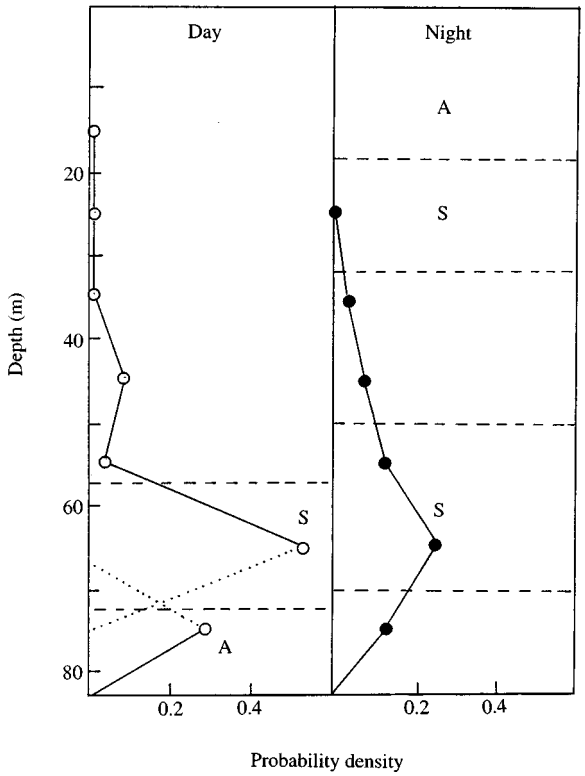


FIG. 7. Normalized distributions of depths of schools of fish derived from echo sounder data measured along the propagation path during day (○) and night (●) respectively. Dotted lines illustrate approximate contributions due to schools of sardines (*S*) and anchovies (*A*) during the day, based on trawling data. The dashed lines represent boundaries of absorption layers attributed to sardines, which were derived from matching transmission loss measurements and computations. Previously published measurements indicate that anchovies are generally found in the top 20 m at night (Holliday, 1972).

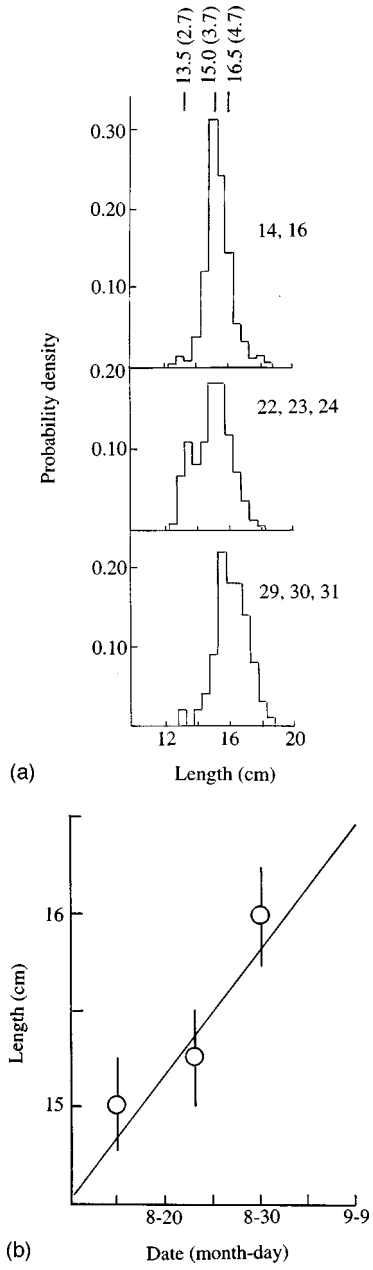


FIG. 8. (a) Normalized distributions (probability per 0.5 cm) of lengths of sardines measured on 14 and 16 August (top), 22, 23 and 24 August (middle), and 29, 30 and 31 August (bottom) at the Modal Lion site. Previously reported average lengths of 2.7, 3.7 and 4.7 year old sardines in the Gulf of Lion (Pichot and Aldebert, 1978) shown at top. (b) Modal lengths of sardines versus data [after 3 point averaging of data shown in (a)]; linearly extrapolated length during Modal Lion: 16.3 cm.

night. Hypothetically, fish at these depths are in both schools and shoals.

Direct sampling of fish during daytime (Liourzou, personal communication) indicated that the layer of schools above the bottom during the day consists of sardines, which occupy a layer which is approximately 20 m above the bottom, and anchovies, which occupy a layer which is approximately 5 m above the bottom. The echo sounder data recorded on board the R.V. ALLIANCE were not sufficiently sensitive to resolve these two layers. The combination of the echo sounder and trawling data suggests that the thickness of each layer is approximately equal to 10 m, as illustrated with the dotted lines drawn in Fig. 7. The effects of sardine schools dominate the echo sounder data, presumably because there were about two times as many sardines as anchovies (Guennegan, 1997), and because the target strength of sardines is about two times larger than the target strength of anchovies.

The R.V. ALLIANCE is a notably quiet ship; consequently, engine noise probably did not have a large effect on measurements of the depths of schools near the bottom during the day. Sequential measurements made in moving and drift modes along the same track during the day (not shown) indicate that the peak of the distribution of the number/km/10 m, when fish are in schools near the bottom, is depressed by less than 5 m, as the ship's speed increases from zero to 9 kts. When fish are at shallower depths this effect is hypothetically more pronounced.

The distribution of fish sizes was sampled in the shaded area shown in Fig. 4 between 14–16, 22–24 and 29–31 August, prior to the commencement of Modal Lion by Liorzou (personal communication). All samples were taken during the day with a trawl that has a mesh size of 1.1 cm. The resultant distributions of fish lengths during these periods are illustrated in Fig. 8(a). The primary modal length of these distributions is approximately 15.5 cm. In addition, a relatively small, secondary peak is discernible during 22–24 August at about 13.5 cm.

Due to the size of the mesh, fish which were smaller than approximately 10 cm were undersampled, and fish which were smaller than 6 cm were essentially not sampled (Liorzou, personal communication). Trawling also preferentially discriminates against larger fish, since they are faster swimmers and can more easily avoid capture (Wardle, 1983; Misund and Algen, 1992; Freon *et al.*, 1993a). The magnitude of the length dependent bias associated with capture avoidance is controlled by the ratio of swimming speed to tow speed. The maximum swimming speed of fish is approximately proportional to the square root of fish length, and is approximately equal to 4 kts for 16 cm long sardines. According to this prescription, the maximum swimming speeds of 18.5 cm long sardines (the largest length which was sampled) is 4.2 kts. Tow speeds are nominally 4 kts. The magnitude of this ratio changes only about 5% between modal and maximum lengths. Consequently, the bias due to length dependent capture avoidance is probably small.

The peak spawning month in the Gulf of Lion is January (Liorzou, personal communication). Consequently, the ages of sardines in September are approximately 8 months, 1 year and 8 months, 2 years and 8 months, etc. Previously reported measurements of the length versus age relationship of sardines in the Gulf of Lion (Pichot and Aldebert, 1978) suggest that the primary peaks in the distribution of sardine lengths, which are shown in Fig. 8(a), corresponds to fish which are 3.7 or 4.7 years old. Despite year to year variability in growth rates, it is instructive to compare results shown in Fig. 8(a) with Pichot and Aldebert's (1978) measurements. The sardines, which contribute to the primary peaks of the distributions shown in Fig. 8(a) (i.e., sardines which are

longer than approximately 14 cm), are 3.7 years old (~15 cm long), 4.7 years old (~16 cm long), 5.7 years old (~16.5 cm long), 6.7 years old (~17 cm long), etc. The much smaller peak, which is discernible during 22–24 August at about 13.5 cm, is consistent with sardines which are approximately 2.7 years old. There are no significant peaks that correspond to younger year classes, viz., sardines, which are approximately 6 cm long (0.7 years old), or approximately 11 cm long (1.7 years old). The absence of a peak at 6 cm is due to mesh size. The absence of a peak at 11 cm, and the relatively small number of 13.5 cm long sardines, may be due to one or a combination of the following. Adult and juvenile sardines may be geographically noncoincident. The number densities of one year old sardines were relatively small.

To estimate the ''effective'' modal length of sardines during Modal Lion, measurements shown in Fig. 8(a) should be extrapolated in time and corrected for the length dependence of the extinction cross section. To minimize the uncertainty in estimates of modal lengths, the results shown in Fig. 8(a) were 3 point averaged. The resultant modal lengths are plotted as a function of time in Fig. 8(b). Based on these results, the modal length on 6 September is estimated to be 16.3 cm. The modal lengths of the primary peaks increased approximately 1 cm per month. This rate of change is much larger than the average growth rate of sardines. The observed increase may be caused, in part, by a higher than average growth rate (which can occur during the summer), and in part, by migration of older, larger fish from near-shore regions to the measurement site. The extinction cross section increases with the square of the radius of the swim bladder (approximately with the square of the fish length). To determine the ''effective'' modal length, measured probability densities were weighted by the square of the length. Based on these calculations, the ''effective'' length was determined to be approximately 0.2 cm larger than the measured modal length, i.e., 16.5 cm on September 6. Finally, the actual ''effective'' modal length may be slightly larger than 16.5 cm due to the length dependent bias associated with capture avoidance.

The modal lengths of anchovies, which were measured during the same periods (not shown) also increased with time, and exhibited a pronounced maximum at 13.5 cm between September 29 and 31. This length corresponds to anchovies, which are 2.2 years old. A smaller peak was also evident at 9.5 cm, which corresponds to anchovies, which are 1.2 years old. The average length of juvenile anchovies during this experiment was about 2 cm (Liorzou, personal communication). Trawling results indicated that the percentage of other fish with swim bladders was small (Guennegan *et al.*, 1997).

Prior to the commencement of Modal Lion the biomass of sardines and anchovies was measured during daytime in the Gulf of Lion using a combination of echo sounder and trawl data. The spatially averaged biomass of sardines and anchovies in the Gulf of Lion in mid-August was estimated to be 7.0 and 2.6 metric tons per km^2, respectively (Guennegan *et al.*, 1997). This implies that there were about two times as many sardines as anchovies in the Gulf of Lion in during this experiment. This calculation of the biomass of sardines was based on an average length of 15.5 cm. The relationship between the weight, W, and the length, L, of sardines in the Gulf of Lion is governed by Eq. (13) (Liorzou, personal communication):

$$W = 0.006L^{3.12}, \tag{13}$$

where W is in grams and L is in cm. Extrapolation of Liorzou's estimate of the spatially averaged biomass to 16.5 cm long sardines [with Eq. (13)] yields 8 metric tons per km^2. Application of Azzali's ''rule'' to the latter estimate implies that the biomass in colonies in the Gulf of Lion, and by inference at the measurement site, was about 32 metric tons per km^2.

The uncertainty in estimates of the biomass, which are derived from echo sounder data, is controlled primarily by avoidance behavior, and by the uncertainty in the target strength of individual sardines during daytime, when they are generally in schools near the bottom (Rose, 1992). Such calculations are based on a limited number of measurements of target strength, which were made at relatively shallow depths. Freon *et al.*'s (1993b) comparison of biomass measurements, which were made during night and day, suggests that the scattering cross section at depths of interest during the day may be 2.2 dB lower than generally employed values (Foote, 1987). This inference suggests that the biomass of adult sardines at the Modal Lion site may have been about 51 tons per km^2. Finally, the actual biomass may have been greater than this estimate as a result of avoidance behavior.

B. Meteorological conditions

Cloud cover was minimal, and the moon was nearly full on 5 and 6 September. A full moon occurred on 8 September. Hypothetically fish can maintain visual contact and form schools at greater depths on moonlit nights than on dark nights, and the percentage of fish in schools on moonlit nights, being more daylike, may be greater than the percentage of fish in schools on dark nights. The mean wind speed was approximately 12 knots from the north near sunrise on 5 September, and approximately 7 kn from 150° near sunset on 5 September and near sunrise on 6 September.

C. Geoacoustics

Approximately one half of the bottom along the measurement track was covered with unconsolidated sediments. The other half consisted of granite. Four cores were taken along the sediment covered portion of the track. The average compressional sound speed, c_P, in the top 2 m was 1560 m/s, and $c_P/c_W = 1.034$, where c_W is the sound speed in water at the ocean-bottom interface. The sound speed in granite was not measured. Its compressional and shear sound speeds are significantly higher than the speed of sound in water. Computations of transmission loss versus depth, which will be discussed in Secs. V and VI, are independent of c_P for $c_P/c_W \geqslant 1.034$. These computations incorporate the directionality of the parametric source, which in this environmental configuration excites predominantly low order modes. These modes are relatively insensitive to c_P when

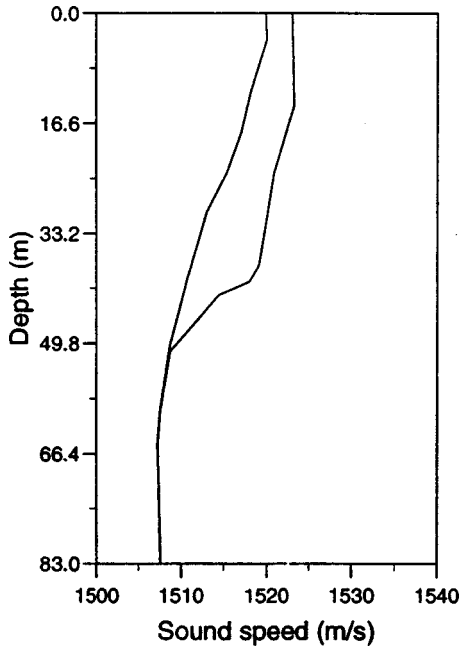


FIG. 9. Average sound speed profiles computed from CTD measurements at the source location (left), and thermistor string measurements at the receiver site (right).

$c_P/c_W \geqslant 1.034$. Computations of transmission loss are also insensitive to the absorptivity in the bottom, primarily because the critical angle, 15°, is relatively small (Brekhovskich, 1960). All computations, which will be presented in Secs. V and VI, incorporate $c_P = 1560$ m/s and an absorption coefficient of zero.

D. Temperature, salinity and sound speed profiles

On 5 and 6 September CTD casts indicated that solar heating during the day and cooling at night had a small effect on measured temperatures at depths down to about 5 m, which is consistent with previously reported measurements of this phenomenon (Price *et al.*, 1986). Below about 5 m the CTD and thermistor chain data revealed essentially no systematic day–night differences in the average profiles, no systematic day–night differences in the variance of the temperature; and no anomalous behavior at twilight. Sound speed profiles derived from time-averaged CTD and thermistor chain data recorded during 5 and 6 September, are illustrated in Fig. 9. The depth dependence of sound speed is controlled primarily by temperature. The average temperature in the mixed layer was approximately 17 °C, and the salinity was 38 psu and independent of depth. The two temperature profiles were qualitatively similar; the mean depth of the thermocline (defined here as the depth at which the temperature gradient is maximum), however, varied between approximately 25 m at the source to 40 m near the receiver. The density of sea water at 20 m was approximately 1.028.

E. Internal waves

Echo sounder measurements at the source, and thermistor data recorded near the receiving array revealed internal waves with dominant periods between 10 and 20 minutes and an amplitude of about 7 m, propagating along the thermocline. The thermistor data indicated that the amplitudes and dominant periods of internal waves were essentially the same during night and day, and that there were no anomalies in the amplitudes of internal waves at sunrise and sunset. Millot and Crepon (1981) provide experimental evidence that the dominant mechanism for generating internal waves in the Gulf of Lion is atmospheric forcing. ''Solitons,'' which are intermittent, tidally generated, nearly monochromatic internal wave packets, were not evident in their data Furthermore, solitons were not evident in thermistor string data recorded during Modal Lion. Such wave packets have been observed at sites where tides are large (Zhou *et al.*, 1991). According to Millot and Crepon (1981) tides are small in the Gulf of Lion

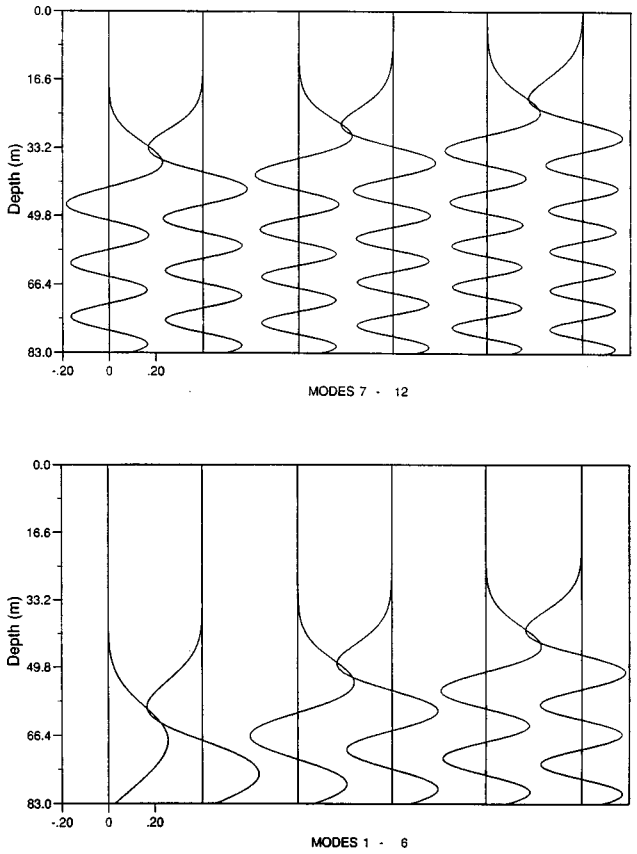


FIG. 10. Eigenfunctions of the first 12 modes at 1.4 kHz computed from the average sound speed profile measured at the source location. The source depth was 78 m.

F. Discussion of environmental effects

Figure 10 illustrates eigenfunctions of the first 12 modes at 1.4 kHz, based on the average sound speed profile derived from CTD measurements recorded on 5 and 6 September, and a bottom sound speed, 1560 m/s, which was derived from core data. One may infer from this figure that a source deployed at a depth of 78 m would excite predominantly modes 2, 3 and 4, and to a lesser extent, modes 1 and 5. The eigenfunctions of modes 2, 3 and 4 at this depth are near maximum. Whereas, the eigenfunctions of modes 5–8 at this depth are relatively small and, as a result are minimally excited. Higher order modes (>8) are attenuated because of the

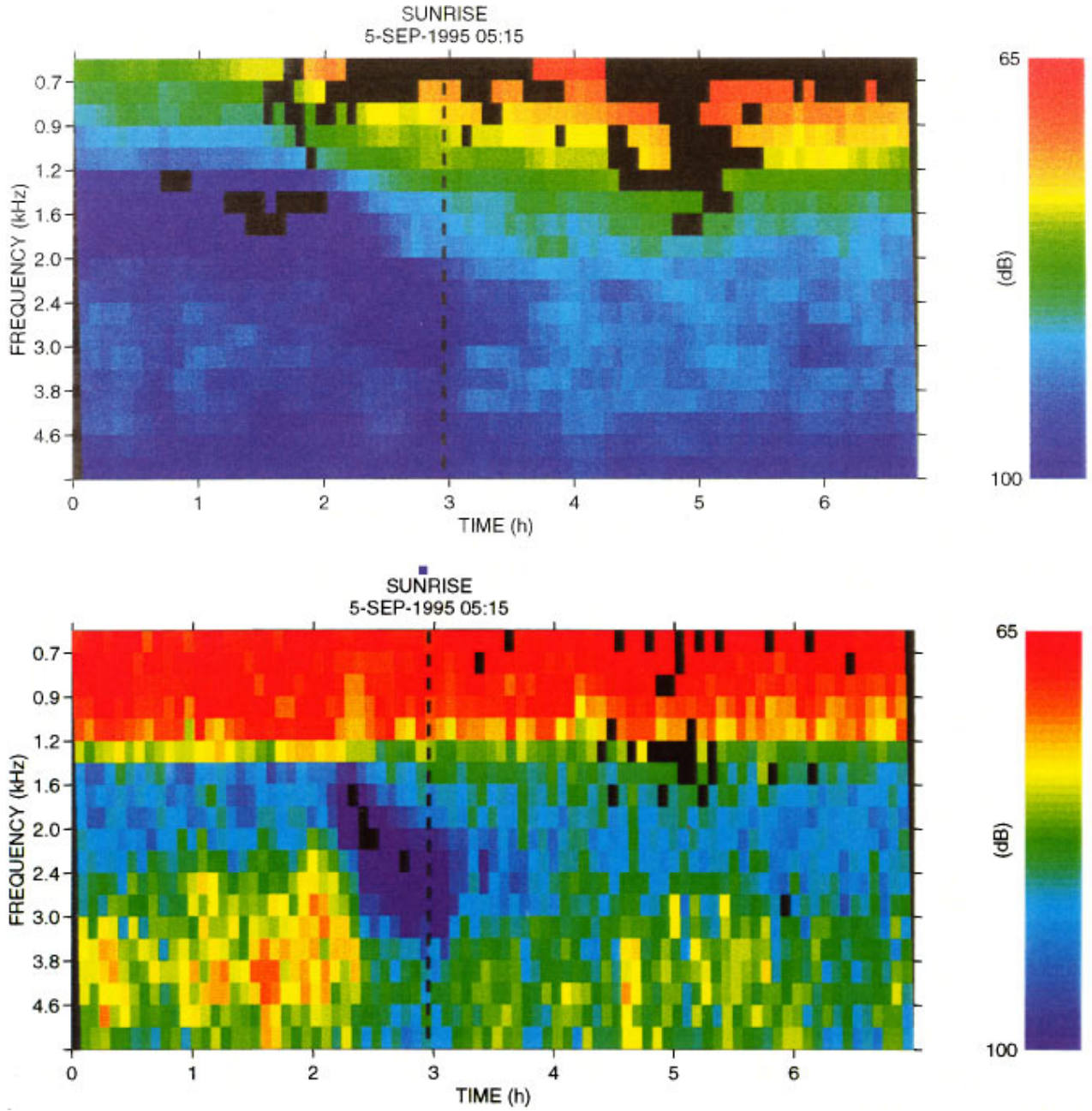


FIG. 11. Transmission loss versus time and frequency on near surface (top) and near bottom hydrophones near sunrise. Low SNR (<2 dB) pixels are black. High absorption at night on near surface hydrophones at ~1.3 kHz is attributed to dispersed sardines at ~20 m; the systematic increase in resonance frequency starting ~1 h before sunrise at ~1.3 kHz and ending at sunrise at ~2.7 kHz, which is evident on near-surface and near-bottom hydrophones, is attributed to dispersed sardines descending to a depth of ~65 m. The rapid shift to lower resonance frequency and lower absorptivity, which occurs ~10 minutes after sunrise, is attributed to the transition of sardines from dispersed mode (where the mean separation between fish, $s > \lambda_0$) to school mode (where $s < \lambda_0$). High absorption at night at ~1.5 kHz, which is evident on near-bottom hydrophones, is attributed to schools at ~60 m. Low SNR during the day at $f < 2$ kHz is due to noise from ships. Low SNR at $f < 0.9$ kHz, which is evident on near-surface hydrophones, is due to low source levels.

directionality of the source. Consequently, in the absence of scattering (mode conversion), modes 2, 3 and 4 may be expected to dominate propagation at 1.4 kHz in this environment. Experimental confirmation of this hypothesis is not possible, as only received energy levels were recorded during this experiment.

This figure may be used to formulate hypotheses about the effects of a thin absorbing layer of fish on acoustic modes. Hypothetically, a thin absorbing layer at 65 m would have a relatively large effect on modes 1, 3 and 4, and minimal effect on modes 2 and 5 at this frequency; whereas a thin absorbing layer at 78 m would have a pronounced effect on modes 2, 3 and 4, and a small effect on modes 1 and 5. Absorbing layers consisting of fish are nominally 10–15 m thick, and would to some extent affect all the modes, in accordance with the depth, thickness and absorptivity of the absorbing layers, and the eigenfunctions of the modes. The effects of internal waves on mode conversion of low to high order modes (scattering of energy into the ''shadow zone'' above the thermocline) will be discussed in Sec. VIII.

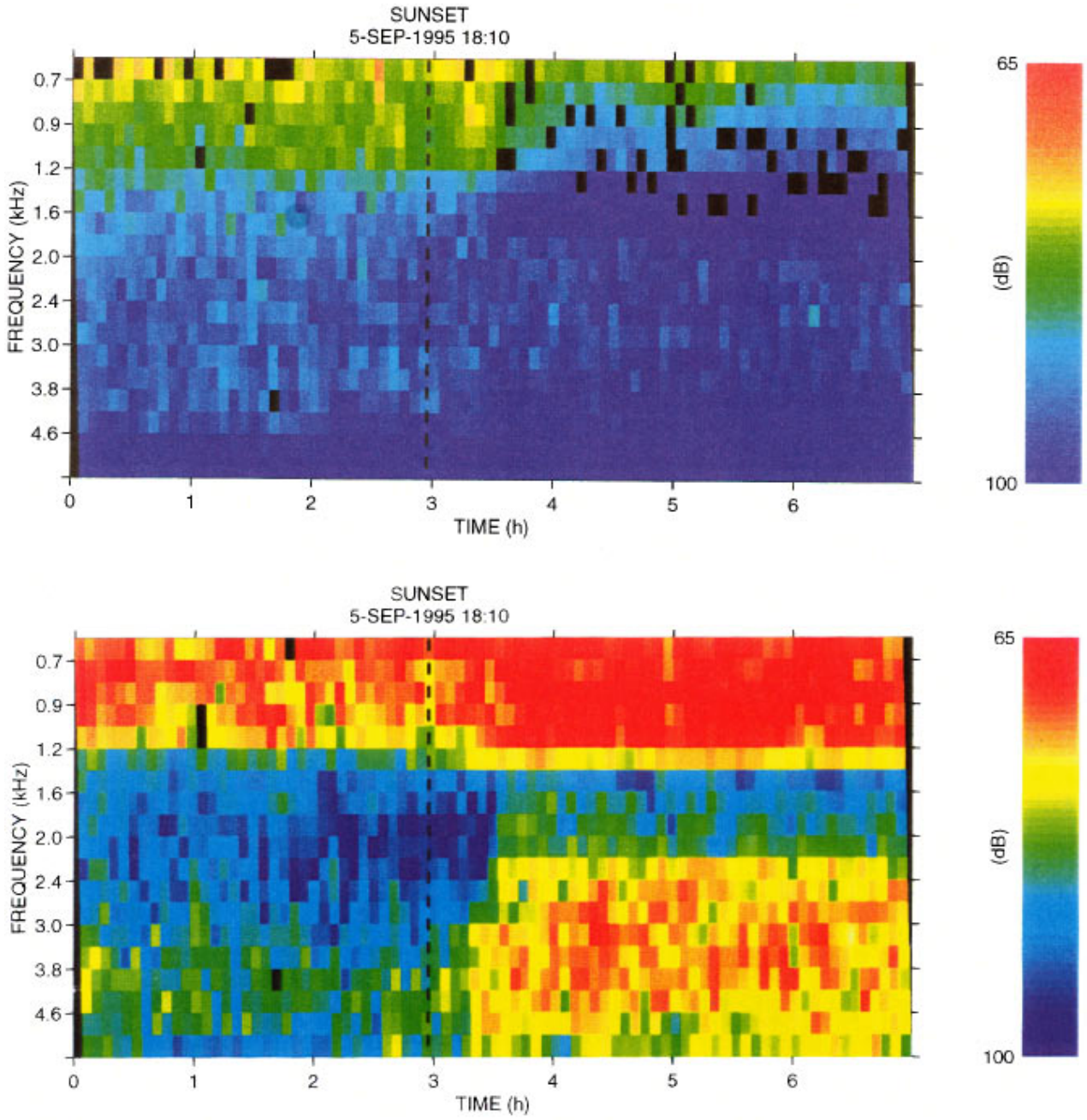


FIG. 12. Transmission loss vs time and frequency on near-surface (top) and near-bottom (bottom) hydrophones near sunset. High absorptivity at ~1.7 kHz during the day, which is evident on near-bottom hydrophones, is attributed to schools at ~65 m. Anomalously higher absorptivity at ~1.7 kHz starting at ~1 h before sunset and ending at ~30 minutes after sunset, which is evident on near-bottom hydrophones, is attributed to increased separation between sardines in schools. Schools slowly ''diffuse'' during sunset twilight (Freon *et al.*, 1996). The absence of a strong frequency dependence during the ascent is attributed to two competing effects that control resonance frequency. As fish ascend, their resonance frequency systematically decrease with time in accordance with Boyle's law. Simultaneously the separation between fish in schools increases with time, causing their resonance frequencies to systematically increase with time. High absorptivity at ~1.5 kHz, which is evident on near-bottom hydrophones starting about 30 minutes after sunset, is attributed to fish schools at ~60 m. The onset of high absorptivity at ~1.3 kHz starting ~30 minutes after sunset, which is evident on near-surface hydrophones, is attributed to the arrival of dispersed sardines at ~20 m.

IV. TEMPORAL EVOLUTION OF RESONANCE FREQUENCIES

Transmission loss was calculated from received signal levels and measured source levels at each transmitted frequency every 5 minutes. Transmission loss measurements made during morning twilight on 5 September are shown in Fig. 11. In this figure the abscissa represents time from beginning of the measurement period, and the ordinate (left) represents the frequency of transmitted signals. The intensity scale is shown on the right. The data in the top figure were averaged over the top three hydrophones; the data in the bottom figure were averaged over the bottom four hydrophones. Averaging minimizes fluctuations and enhances systematic changes in signal level versus time and frequency. The variance is relatively low at frequencies below approximately 2 kHz, where effects due to internal waves appear to be small.

The data shown in Fig. 11 exhibit the following features:

relatively high losses in signal level (about 18 dB) at about 1.3 kHz at night (evident in the top figure), and a strong systematic frequency dependence of transmission loss during twilight (evident in both figures). This is consistent with the presence of dispersed sardines at night at a depth of 20 m, where the resonance frequency is approximately 1.3 kHz, followed by their twilight descent from 20 m to 65 m, which starts approximately 85 minutes before sunrise, and is completed at approximately 5 minutes after sunrise, where the resonance frequency is approximately 2.8 kHz. This two-step (downward descend followed by school formation) process during sunrise is consistent with Freon's (1996) observations of the structure of schools of sardines during the sunrise descent. The observed temporal evolution of the resonance frequency of the absorption line is consistent with echo sounder data recorded during this period of time, shown in Fig. 6.

In Fig. 11 low signal to noise pixels are black. During the sunrise descent signals between 1.3 and 2.7 kHz are preferentially absorbed, whereas during daytime, noise due to fishing vessels is intermittently high below 2 kHz.

Measurements on near-bottom hydrophones (Fig. 11, bottom) are dominated by the first few modes. These modes are sensitive to absorbing layers below the thermocline, and are consequently not affected by the layer of dispersed fish at 20 m. On the other hand, measurements on near-surface hydrophones (Fig. 11, top), are dominated by high order modes, which are nominally equally sensitive to absorbing layers throughout the water column, including the layer of dispersed fish at 20 m. Highest transmission losses (110 dB at 2.7 kHz) were observed on near-bottom hydrophones approximately 5 minutes after sunrise, when all (or most) of the fish reached 65 m, in accordance with echo sounder data. Approximately 10 mins after sunrise absorption losses abruptly diminish and the resonant frequency is lowered to approximately 1.7 kHz, consistent with the formation of schools, which was also observed in the echo sounder data. At night an absorption line is also evident on near bottom hydrophones at about 1.5 kHz, hypothetically due to schools of sardines near the bottom. Measurements made on 6 September around sunrise exhibit essentially the same pattern shown in Fig. 11.

For comparison, transmission loss measurements taken during evening twilight are shown in Fig. 12. On near-bottom hydrophones transmission losses are highest at about 1.7 kHz during daytime, and 1.5 kHz during nighttime. In contrast to the data at sunrise, the data at sunset are characterized by the absence of a deep, strongly frequency dependent null during the ascent at sunset. This frequency dependence is consistent with the temporal evolution of the structure of schools of sardines during sunset, observed by Freon *et al.* (1996). During sunset sardines simultaneously rise, which causes f_0 to systematically decrease with time, and slowly ''diffuse'' or disperse from schools, which causes f_0 to systematically increase with time.

V. THE TRANSMISSION LOSS MODEL

To permit simulation of these observations the coupled normal mode model, C-SNAP (Ferla *et al.*, 1993), was modified by Ferla (Diachok and Ferla, 1996) to include the beamwidth of the source, and absorbing layers in the water column. The resultant program permits weighting the energy of modes emitted by the source in accordance with a Gaussian weighting function (to account for the beamwidth of the source). Absorbing layers of selectable depth, *d*, thickness, *t*, and attenuation coefficient, permit attenuation of each mode proportional to the energy of the mode at the depth of the absorbing layers.

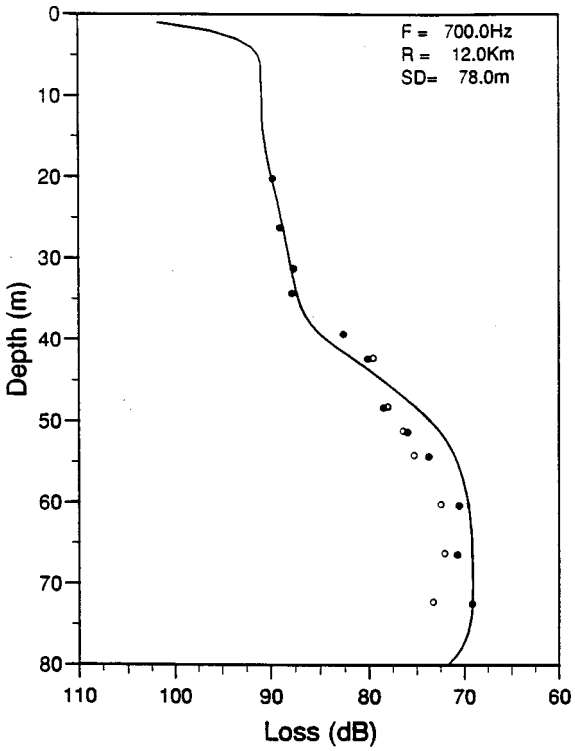


FIG. 13. Transmission loss measurements at night (●) and during day (○) at 0.7 kHz versus depth. Note that there are no data (SNR<5 dB) during day on hydrophones above the thermocline. Theoretical computations are for α equal to 0 dB/km.

Transmission loss computations presented in this paper are based on average oceanic sound speeds calculated from the CTD and thermistor chain measurements made on opposite ends of the propagation path, and a sound speed in the bottom of 1560 m/s in accordance with measurements made on four cores. Sound speed profiles were averaged over two days. Sound speed profiles between the source and receiver were interpolated. Results of transmission loss computations at 0.7 kHz, which exclude absorbing layers in the water column, are shown in Fig. 13. These computations, which represent an incoherent sum of modes, yield an excellent match (± 2 dB) to time-averaged transmission loss measurements (also shown in Fig. 13). These data were averaged over one hour periods during day and night. The agreement between computations and measurements is slightly better at night. Coherent summation of modes yields a much stronger variability in the depth dependence (due to modal interference), which is not evident in the data. The loss of intermodal phase coherence may be caused by mode coupling due to bottom roughness, internal waves or larger scale spatial variability of sound speed structure along the measurement track.

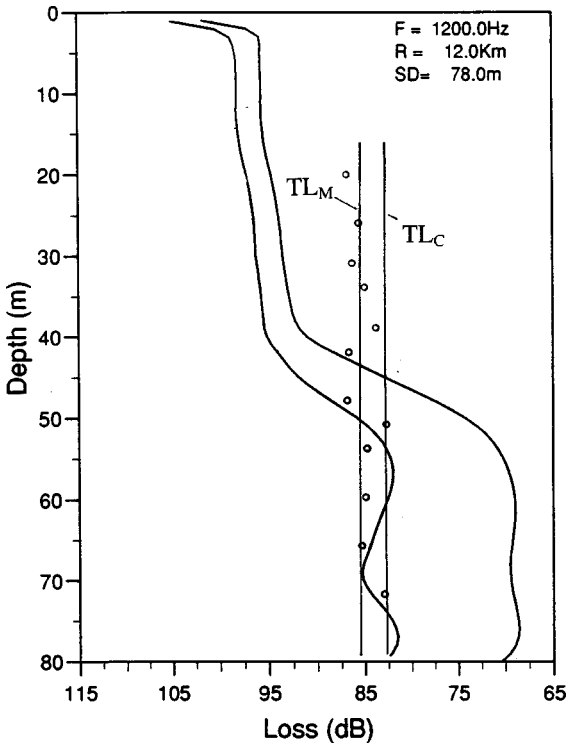


FIG. 14. Transmission loss measurements (○) during day at 1.2 kHz versus depth. Also shown are computations without absorbing layers (right), and with an absorbing layer of fish schools (left). The absorption coefficient, α equals 2.4 dB/km; the depth and thickness of the layer equal 65 m and 15 m, respectively. TL_M is the average measured value, and TL_C is the average corrected (for absorptivity) value of transmission loss above the thermocline.

The good agreement between transmission loss measurements and computations at 0.7 kHz at near bottom depths suggests that (1) the assumed sound speed model is valid, and (2) the source level of the parametric source was not significantly affected by environmental effects in the source generating region, such as internal waves, bottom interaction, bubbles, suspended sediments and fish. The small discrepancy evident during the day is probably a result of increased α during the day (due to a larger number of schools near the bottom during the day).

Conceivably schools may occasionally pass through the source generation region and affect source levels. In principle energy levels at the difference frequency may be increased due to the nonlinear properties of swim bladders, and decreased due to increased absorptivity at pump frequencies (Kustov *et al.*, 1982; Druzhinin *et al.*, 1996). Ching and Weston (1975) attributed high level ($\sim\pm 10$ dB) fluctuations in received levels during daytime to the passage of fish schools through the near field of their sources. The standard deviations of transmission loss measurements recorded during day and night during Modal Lion were approximately equal (± 3 dB), suggesting that this effect was small during this experiment.

The signal to noise ratio of measurements made on hydrophones above the thermocline during the day at 0.7 kHz was too low to permit comparison with computations. The good agreement between measurements made on near surface hydrophones at night and computations is fortuitous, since these computations disregard the effects of forward scattered energy above the thermocline, which occurs during both day and night, and absorptivity due to dispersed fish above the thermocline, which occurs at night.

VI. EFFECTS OF ABSORBING LAYERS

Transmission loss measurements at 1.2 kHz, which were made during the day, are illustrated in Fig. 14. These measurements represent an average over a one hour period. At this frequency the signal to noise ratio is sufficiently high on all hydrophones to permit comparison of measurements made during daytime at depths above the thermocline with theoretical computations. These computations illustrate the effect of an absorption layer which has a thickness of 15 m and a depth of 65 m. Computations are provided for α equal to zero and 2.4 dB/km. At this frequency measurements of transmission loss below the thermocline exceed computations by about 15 dB.

Above the thermocline measured energy levels at 1.2 kHz during daytime exceed computations by about 10 dB. The anomalous high energy level above the thermocline is attributed to forward scattering by bottom roughness (Rouseff and Ewart, 1995) or internal waves (Schneider, 1990; Thiele, 1990; Sellschopp, 1990). Properties of forward scattered energy above the thermocline will be discussed in Sec. VIII.

The effects of two absorption layers on transmission loss at 1.2 kHz at night are considered in Fig. 15. The measurements, which are shown in this figure, were averaged over a period of one hour at night. Computations are provided for the nighttime scenario, which consists of a 15 m thick layer of dispersed fish at 25 m, and a 20 m thick layer of fish in schools at 60 m. Computations for $\alpha=0$ are shown for reference. These computations disregard forward scattered energy above the thermocline (which is illustrated in Fig. 14); hence, the resultant estimate of α of the layer of dispersed sardines at 25 m should be considered a lower bound. Actual losses will be shown (in Sec. X) to be higher. Matching of the measured and computed transmission loss versus depth provides estimates of the average depth and thickness of absorption layers, and a lower bound on α.

Computations, which incorporate other values of d, are illustrated in Fig. 16. These calculations suggest that the average depth of this absorbing layer is 25 ± 3 m. This is consistent with the magnitude of $d(20\pm 3\,\mathrm{m})$, which may be derived from the measurements of $f_0(1.3\,\mathrm{kHz})$, $r_0(0.79\,\mathrm{cm})$, and $\epsilon(12)$, and Eq. (4).

Another potential source of error is due to mode coupling, which could conceivably bias estimated depths of absorbing layers through the following process: (1) low to high order mode conversion, (2) absorption of high order modes by the near surface absorbing layer, and (3) high to low order mode conversion. This phenomenon may also conceivably bias estimates of the magnitude of the absorption coefficient due to schools (near the bottom) at night. Theoretical investigation of this hypothesis awaits development of a model that incorporates both mode conversion and absorbing layers.

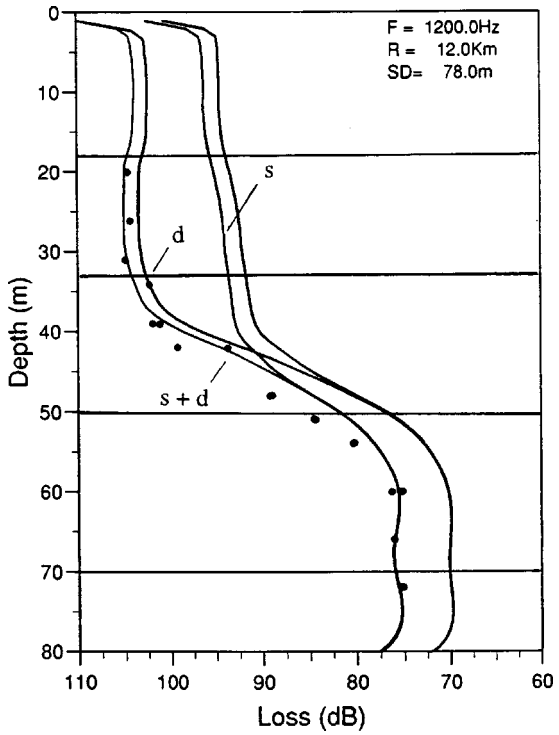


FIG. 15. Transmission loss measurements (●) at night at 1.2 kHz vs depth without absorbing layers (right), and computations with absorbing layers of dispersed fish (*d*) at 25 m (layer thickness: 15 m), fish schools (*s*) at 60 m (layer thickness: 20 m) and both layers (*s*+*d*). The absorption coefficient, α, equals 1.25 dB/km in the near surface layer and 5.0 dB/km in the near bottom layer.

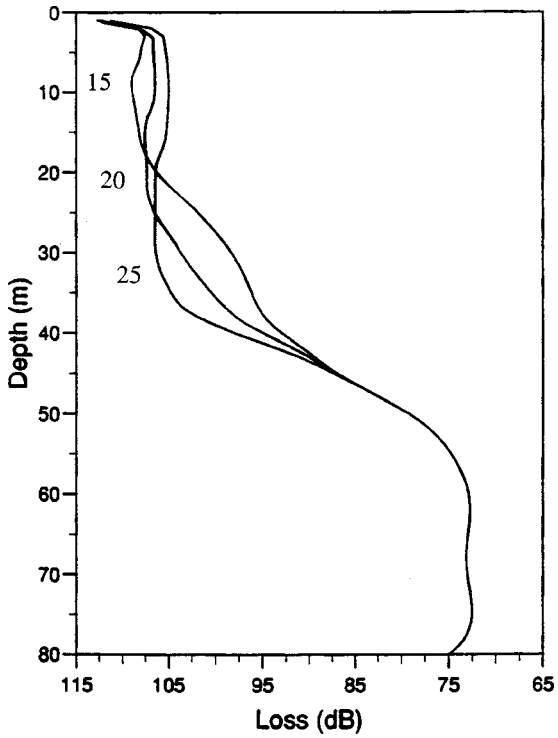


FIG. 16. Computations of the effect of an absorption layer on transmission loss at 1.2 kHz centered at depths of 15, 20 and 25 m. The absorption coefficient of the layer is 5 dB/km, and its thickness is 15 m. All computations include an absorption layer centered at 60 m, with a thickness of 20 m, and an absorption coefficient of 1 dB/km.

Transmission loss measurements during night and day are consistent with calculations that include an absorbing layer at a depth of 62.5 ± 3 m. This result is consistent with echo sounder data, which indicated that $d \approx 65 \pm 3$ m during day and at night. This result is also consistent with trawling measurements, which revealed that the average depth of sardines during the day is about 20 m above the bottom, i.e., at $d \approx 63$ m at the Modal Lion site.

VII. ABSORPTIVITY DUE TO FISH SCHOOLS

Estimates of α, which are due to the layer of schools of sardines at night, are shown in Fig. 17. These estimates were generated by matching measurements and computations of transmission loss at depths below the thermocline. Computations included an absorbing layer, which was 60 m deep and 20 m thick. The magnitudes of α shown in Fig. 17 provided the best fit to the data. The resonance frequency of schools of sardines at night, f_0', is 1.5 ± 0.1 kHz. At frequencies above 2.4 kHz departures from geometrical spreading on near bottom hydrophones at night were small. This suggests that (1) effects due to the layer of schools of sardines in this frequency range are small, and (2) effects of scattering losses (due to bottom roughness and/or internal waves) on low order modes in this frequency range (2.4–4.0 kHz) are small (<2 dB).

Estimates of α as a function of frequency during the day are also shown in Fig. 17 [previously reported estimates of α versus frequency during daytime (Diachok, 1996; Diachok and Ferla, 1996) disregarded effects of α on the forward scattered component, and are slightly lower at near resonance frequencies]. Measurements of transmission loss during the day are characterized by a minimum at 66 m over the frequency range 1.2–1.8 kHz. The resonance frequency of schools of sardines during the day is 1.7 ± 0.1 kHz, which is slightly higher than f_0' of schools of sardines at night (1.5 ± 0.1 kHz). Also evident in the daytime data is a small peak at 2.7 kHz, which is hypothetically due to dispersed sardines.

VIII. FORWARD SCATTERED ENERGY ABOVE THE THERMOCLINE

Previously reported measurements of transmission loss from a source below the thermocline to a receiver above the

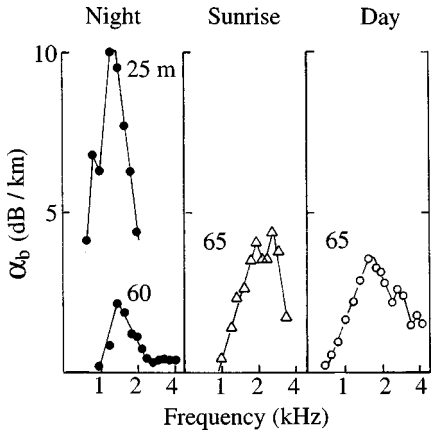


FIG. 17. Absorption coefficients of layers of sardines versus frequency and depth at night (●), day (○), and sunrise (△).

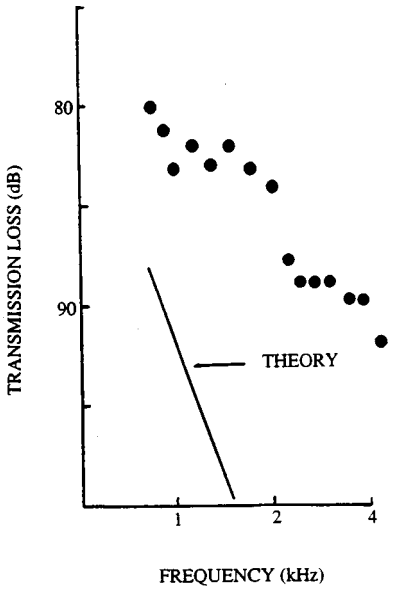


FIG. 18. Theoretical computations and measurements of transmission loss versus frequency (corrected for absorptivity) on top three hydrophones during daytime.

thermocline at another (~110 m) shallow water site indicate that energy levels above the thermocline at a range of about 12 km were about 15 dB higher than deterministic adiabatic mode predictions. Schneider (1990), Thiele (1990) and Sellschopp (1990) attributed these anomalous high energy levels to scattering by internal waves. Computations due to Rouseff and Ewart (1995) suggest that this effect may be caused by bottom roughness.

During Modal Lion measured transmission loss on near surface hydrophones during the day on 6 September was about 86 dB (TL_M in Fig. 15). This energy level is about 10 dB higher than theoretical computations at 1.2 kHz. If there were no absorbing layer at 65 m, then forward scattered energy level would be 2–3 dB higher, and the corresponding transmission loss (TL_C in Fig. 14) would be about 83 dB. The increase in signal level may be estimated from the equation

$$\Delta = 12\alpha t/83, \tag{14}$$

where Δ (in dB) is the increase in signal level, α (in dB/km) is the absorptivity in the layer of schools during the day, t is the thickness (in m) of the layer, 12 km is the range and 83 m is the depth of the water column. Similar computations at frequencies between 0.8 and 4.0 kHz, illustrated in Fig. 18, indicate that the energy in forward scattered modes decreases systematically with frequency. The most likely causes of the forward scattered energy are bottom roughness and internal waves. The spectrum of energy scattered into high order modes is controlled by the process of mode conversion, and by attenuation of high order modes by bottom and surface roughness and subsurface bubble clouds. These attenuation mechanisms increase with frequency in this frequency range.

The hypothesis, that forward scattered energy may be caused by sound speed anomalies associated with high concentrations of schools near the bottom during daytime, seems unlikely. Forward scattered levels are highest at the lowest frequencies and decrease systematically between 0.8 and 4.0 kHz. This observation is in sharp contrast with measured absorption losses, which peak at the resonance frequency of schools, 1.7 kHz.

IX. ABSORPTIVITY DUE TO DISPERSED FISH VERSUS FREQUENCY AT NIGHT

Since there is no systematic diel variability in the sound speed profiles and in the average amplitudes of internal waves, and bottom roughness is time invariant (over the duration of the experiment), the forward scattered signal may be assumed to be the same during day and night. Consequently, differences between measurements made during night and day on near-surface hydrophones are probably a more accurate indicator of the effects of absorption losses due to the dispersed layer at 20 m than differences between measurements made at night and theoretical computations. This assumption, which is consistent with Ching and Weston's (1971) analysis of their data, disregards diurnal changes in the effects of fish schools on high order modes. The difference between measurements of transmission loss on near surface hydrophones (Δ_T) during night (TL_N) and day (TL_D) may be expressed as

$$\Delta_T = TL_N - TL_D = \Delta_N(60) + \Delta_N(20) - \Delta_D(65), \tag{15}$$

where $\Delta_N(60)$ is the loss in signal level of high order modes due to a layer of schools at 60 m at night, $\Delta_N(20)$ is the loss in signal level due to a layer of dispersed sardines at 20 m at night, and $\Delta_D(65\,\mathrm{m})$ is the loss in signal level due to a layer of schools of sardines during the day. These computations employ estimates of α shown in Fig. 17, and Eq. (14). Rearranging Eq. (15) permits computation of

$$\Delta_N(20) = \Delta_T + \Delta_D(65) - \Delta_N(60\,\mathrm{m}) \tag{16}$$

from which a more reliable estimate of the absorptivity in the layer of dispersed fish at night may be computed. Results of these computations, as a function of frequency, are shown in Fig. 17. The resonance frequency of dispersed sardines in this layer is 1.3 ± 0.1 kHz. This frequency is consistent with the theoretical value of f_0 due to adult sardines at 25 m, viz., 1.5 ± 0.3 kHz.

The uncertainty in the theoretical value of f_0 is controlled by uncertainties in the estimated values of r_0 and ϵ. The data shown in Fig. 1 suggest that the uncertainties in the extrapolated values of r_0 and ϵ of 16.5 cm long sardines are approximately $\pm 10\%$, and $\pm 20\%$, respectively.

The magnitudes of r_0 and ϵ of 16.5 cm long sardines in the Gulf of Lion and the Bay of Biscay may be different. The relationship between L and r_0 depends to some extent on the percent oil content (Blaxter and Batty, 1990). Measurements of the swim bladders of adult herring (Brawn, 1969; Blaxter and Batty, 1990) suggest that the magnitudes of r_0 are inversely correlated with percent oil content. Modal values of r_0 for 16.5 ± 1.5 cm long herring with high (~17%) oil content were about 0.88 times smaller than r_0 for herring with low (~1%) oil content. This parameter increases systematically with L, varies cyclically with season, is site dependent, and varies randomly from year to year (Love, 1970). The oil content of adult sardines in the western Mediterranean (be-

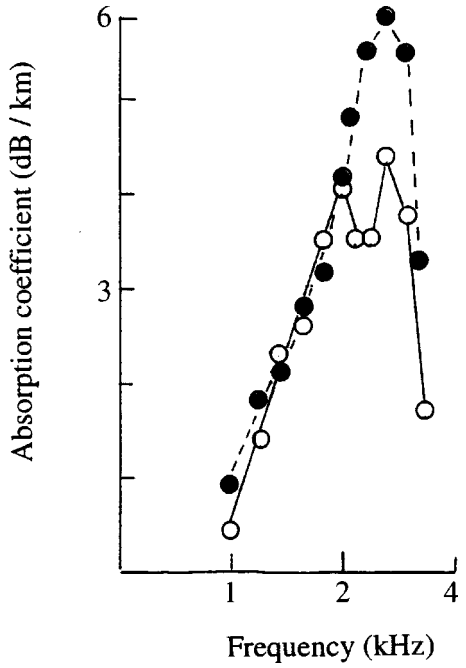


FIG. 19. Absorption coefficients versus frequency at sunrise on 5 September (●) and 6 September (○).

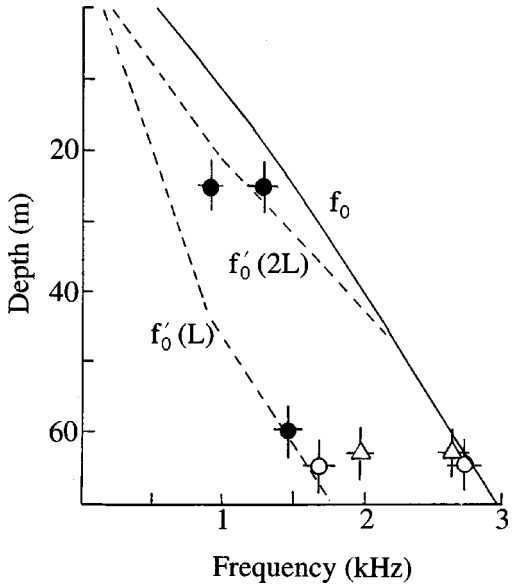


FIG. 20. Measured resonance frequencies of sardines at night (●), day (○) and sunrise (△); and theoretical computations of the resonance frequencies of dispersed sardines (f_0), and sardines in schools (f_0') for s equal to L and $2L$, and N equal to 10^4.

tween 1953 and 1955) was maximum (~8%) in mid-summer and minimum (~1%) during mid-winter (Herrera and Munoz, 1957). For comparison, the oil content of adult sardines at a site in the Atlantic (between 1936 and 1938) was maximum in December (~17%) and minimum in April (~6%) (Hickling, 1945). The experiment, described in this paper, was conducted in September. Scalabrin's measurements in the Bay of Biscay were made in May. The percent oil was not measured at either site. These observations suggest that magnitudes of r_0 in the Gulf of Lion and Bay of Biscay may differ as much as ±12% due to this effect. Consequently, the total uncertainty in the calculated value of f_0 is approximately ±20%.

Previously reported measurements due to Pichot and Aldebert (1978) may be employed to calculate the average lengths of year classes in the Gulf of Lion. Based on their results, 2.7 year old sardines are approximately 13.5 cm long, 3.7 year old sardines are 15 cm long, 4.7 year old sardines are 16 cm long, 5.7 year old sardines are 16.5 cm long, and 6.7 year old sardines are 17 cm long. Based on results shown in Fig. 1 and Eq. (4), these lengths correspond to the following resonance frequencies at a depth of 20 m: 1.5, 1.4, 1.3, 1.3, 1.2 kHz. The peak of the measured absorption line (~1.3 kHz) corresponds to sardines, which are approximately 3.7 and 4.7 years old. According to the same calculations, sardines which are 0.7 years old (0.6 cm) resonate at a much higher frequency (about 4 kHz at this depth), and do not contribute to this absorption line. Analysis of absorption coefficients due to juvenile sardines is beyond the scope of this paper.

It is unlikely that the line at 1.3 kHz is due to anchovies. In a region where both sardines and anchovies were present Barange *et al.* (1996) found that anchovies occupy a layer centered at about 9 m at night. The resonance frequency of dispersed 13.5 cm long anchovies at this depth is 1.1 kHz (assuming that r_0 equals 0.61 cm, ϵ equals 8 and x equals 1.18). Matching measurements and calculations of transmission loss at frequencies below 2 kHz is consistent with an absorbing layer at about 25 m at night. Conceivably dispersed anchovies may contribute to the level and width of this line. However, the contribution due to anchovies is probably small, because the number of anchovies is relatively small (1/2) compared to the number of sardines, and the magnitude of σ_E of anchovies is small (~1/2) compared to the magnitude of σ_E of sardines (at the same depth).

X. ABSORPTIVITY OF DISPERSED SARDINES DURING DAWN AND DAY

Estimates of the absorptivity of dispersed sardines at sunrise on 6 September are also shown in Fig. 17. These estimates were derived by matching transmission loss measurements, which were averaged over the last three samples prior to sunrise, with theoretical calculations. The latter assumed a 15 m thick absorbing layer at an average depth of 65 m. The absorption loss at a depth of about 65 m at a frequency of about 2.7 kHz at sunrise was approximately 26 dB. The estimated absorption coefficient at this frequency was 4.3 dB/km. The estimated uncertainty in f_0 is ±0.3 kHz. The depth was assumed to be at the depth of sardines in school during the day. The corresponding theoretical value of f_0 (at a depth of 65 m), 2.8±0.6 kHz, is in good agreement with measurements. The smaller peak, at 2.0 kHz, which will be discussed in the next section, is attributed to schools of sardines.

For comparison, similar measurements, which were made at sunrise of 5 September, are shown in Fig. 19. The resultant resonance frequency, 2.7 kHz, was the same as on 6 September; and the absorption coefficient at this frequency, 5.8 dB/km, was about 35% larger than on 5 September.

According to Azzali (1985) about 40% of pelagic fish during daytime are either dispersed or in shoals. The daytime spectrum at 65 m, which is shown in Fig. 17, exhibits a large peak at 1.7 kHz, and a smaller peak at 2.7 kHz. The latter frequency is nearly equal to the calculated resonance frequency of dispersed sardines at 65 m, 2.8 ± 0.6 kHz. These data suggest that a large percentage of sardines were either dispersed or in shoals during daytime during this experiment.

The resonance frequencies of dispersed sardines recorded during night, sunrise and day are plotted as a function of depth in Fig. 20. Also shown are theoretical computations based on Eq. (4). Experimental and calculated values are in good agreement.

It is unlikely that the lines at 1.7 and 2.7 kHz are due to dispersed anchovies. These fish occupy a depth of about 78 m at sunrise and during daytime. The resonance frequency of dispersed anchovies at this depth is 3.8 kHz. The small peak evident at this frequency in the data shown in Fig. 17 may not be significant.

XI. RESONANCE FREQUENCIES OF SARDINES IN SCHOOLS

Measurements of the resonance frequencies of sardines in schools are also shown in Fig. 20. The resonance frequencies of sardines in school, f_0', are substantially smaller than corresponding theoretical and measured values of f_0 of dispersed sardines during day (1.7 kHz) and night (1.5 kHz) at the same depth. The difference between measurements of f_0' and $f_0 (\sim 1$ kHz) at ~ 65 m is significantly greater than the combined effects of the uncertainty in the measurements of f_0' and f_0. The ratio, f_0'/f_0, is equal to approximately 0.6 during day.

Figure 20 also shows computations of f_0', based on Eq. (12), as a function of depth and frequency for N equal to 10^4 and s equal to L and $2L$. Calculations for N equal to 0.5×10^4 and s equal to $0.8L$ are essentially identical.

According to Freon *et al.*'s (1996) observations, the volume of sardine schools at night is approximately 0.6 times the volume of schools during day. Smaller volumes imply smaller values of N at night. The number of schools per km during the Modal Lion experiment during night and day were approximately equal. If it is assumed that the number of sardines in schools during the day is about two times the number in schools at night, then N during the day is about two times N at night. The experimental observation of the frequency shift associated with schools during night is consistent with theoretical computations [according to Eq. (12)], provided that both N and s are smaller at night. For example, a combination of nighttime values of N equal 0.6 times daytime values, and nighttime values of s equal 0.8 times daytime values, is consistent with the measured frequency shift. This implies that sardines in schools at about 65 m were slightly closer to each other at night ($\sim 0.8L$) than during daytime ($\sim 0.8L$). Both are comparable to measurements of s of sardines, $0.6L$, which were made in a laboratory (Cullen *et al.*, 1963).

It is unlikely that the line at 1.7 kHz during daytime is due to schools of anchovies at 78 m. The modal horizontal and vertical dimensions of schools of 13.7 cm long anchovies, which were measured by Scalabrin (1997), were 12 and 1.3 m, respectively. The effective radius and eccentricity of such schools are 2.9 m and 9.2, respectively. The corresponding magnitudes of N and β (assuming that s equals L) are 4×10^4 anchovies per school and 3.5×10^{-4}, respectively. According to Eq. (12), the magnitude of f_0'/f_0 is 0.31, and f_0' is 1.2 kHz. An absorption line at this frequency is not evident in the data. However, absorptivity due to schools of anchovies may contribute to the level and width of the line at 1.7 kHz. Their contribution is probably small, since the relative number of anchovies is small ($\sim 1/2$) compared to the number of sardines, and the relative magnitude of σ_E' of anchovies at 78 m is small ($\sim 1/2$) compared to the magnitude of σ_E' of sardines at 65 m.

The slight peak at 0.9 kHz associated with the absorbing layer at ~ 25 m at night (shown in Fig. 17), may be due to schools of 16.5 cm long sardines. Presumably schools of sardines occupy the same depth as sardines which are in shoals or dispersed. Calculated values of f_0' are also provided in Fig. 20 for s equal to $2L$ and N equal to 10^4. The magnitude of f_0'/f_0 that corresponds to this peak is 0.69 (assuming that f_0 equals 1.3 kHz). This ratio is equal to the calculated magnitude of this ratio at 20 m for s equal to $2L$. This estimate is within experimental bounds. Freon (1996) observed that the average separation between sardines in schools near the surface at night at his measurement site was about $3L$.

The absorption coefficient versus frequency during sunrise on 6 September, which is shown in Fig. 17, exhibits a peak at 2.0 kHz. This result is consistent with s equal to about $1.5L$, and N equal to 10^4. There are no secondary peaks evident in measurements of the absorption coefficient versus frequency on 5 September, which suggests that the percentage of fish in schools during that sunrise was small.

The measurements of peak resonance frequencies, which are shown in Fig. 20, suggest that absorptivity due to sardines manifests itself in pairs of absorption lines per layer of sardines. These lines correspond to the resonance frequencies of an ensemble of individual sardines, and the resonance frequencies of an ensemble of ''nuclei'' of schools. This observation is consistent with Misund's (1991) and Aoki and Inagaki's (1988) observations of the distributions of s in schools and shoals.

XII. Q_0 OF DISPERSED SARDINES

According to Eqs. (5) and (6), knowledge of the magnitude of Q_0 of individual fish is required to estimate number density (which will be considered in Sec. XIII). The spectral width of the absorption line associated with dispersed sardines may be approximated with the following equation, which is a modified form of an equation due to Weston (1967):

$$1/Q_T^2 = 1/Q_Z^2 + 1/Q_L^2 + 1/Q_0^2, \tag{17}$$

where Q_Z is associated with the distribution of fish in depth, Q_L is associated with the distribution of fish in length, and Q_0 is the inherent Q of a single fish. Other possible causes of spectral spreading will be assumed to be small.

Experimental measurements indicate that physical contact between the swim bladder and fish flesh has a large effect on Q_0. Laboratory measurements by Baltzer and Pickwell (1970) indicate that Q_0 of swim bladders extracted from anchovies was ~21 at a depth of 6 m, whereas Q_0 of anchovies at the same depth was ~4.5. Sundness and Sands's (1975) measurement of Q_0 of char at 6 m was ~4.9. Lovik and Hovem's (1979) measurements of Q_0 of sprat, and herring were ~2.98 and 2.2, respectively, at this depth. Disregarding the measurements made on herring, which may be anomalous, the average value of Q_0 equals ~4.1 at this depth. This value is virtually identical to the average value of Q_0, which was determined from a large number of measurements at sea by Andreeva (1964). The dependence of percent oil content on Q_0 is unknown; and the extent to which these measurements may have been affected by swim bladder compression by internal organs is also unknown.

The magnitude of Q_0 increases with depth. According to Andreeva (1964), the magnitude of Q_0 at 20 m is 5. The thickness of the absorbing layer at this depth was estimated (in Sec. VII) to be 15 m. For dispersed sardines distributed over 15 m at an average depth of 20 m, $Q_Z \approx 2.2$. For the distribution of lengths shown in Fig. 8, $Q_L \approx 9$. Assuming that Q_Z equals 2.2, Q_L equals 9, and Q_0 equals 5.0, Eq. (17) yields Q_T equal to 2.0. The measured value of Q_T of this line is about 1.3. The width of this line may be broadened by effects due to dispersed anchovies (f_0 equals 1.1 kHz at 7 m), and schools of sardines and anchovies.

According to Andreeva (1964), the magnitude of Q_0 at 65 m is 7. If the thickness of the absorbing layer at 65 m is assumed to be 15 m, then Q_Z equals 5.4. Assuming these parameters and Q_L equal are to 9 in Eq. (17) yields Q_T equal to 3.9. The measured value of Q_T is 1.9. The width of this line may also be broadened by effects, which are due to schools of sardines and anchovies.

XIII. ESTIMATION OF NUMBER DENSITY

The objective of this section is to present a heuristic model that will illustrate how number density and biomass may be derived from absorptivity measurements. Resultant estimates should be considered ''provisional,'' since uncertainties in some of the input parameters are difficult to quantify; e.g., the contribution due to anchovies is unknown. According to Eqs. (5) and (6), the absorption coefficient at f_0 due to a large number of identical fish may be described by:

$$\alpha = n_E r Q_0 \lambda_0, \tag{18}$$

where n_E is the effective number density of acoustically identical sardines. The percentage of sardines, which are acoustically identical, can be estimated from the percentage of sardines that occupy a depth interval in which Q_Z is larger than Q_0. Based on this assumption,

$$n_E \approx n(\Delta t/t), \tag{19}$$

where n is the total number density of adult sardines, Δt is the depth interval over which sardines are assumed to be identical, and t is the thickness of the absorbing layer. It will be assumed that sardines are identical in depth intervals in which $Q_Z \approx Q_0$. This condition is satisfied at night when Δt equals 7 m, and at sunrise when Δt equals 12 m. Substituting Eq. (19) into (18) yields:

$$\alpha = n r Q_0 \lambda_0 (\Delta t/t). \tag{20}$$

In the following computations the magnitude of r will be based on measurements of $r_0(0.0079\,\text{m})$, extrapolated to 20 and 62 m, to be consistent with measurements of f_0. For the dispersed layer of sardines at night, f_0 equals 1.3 kHz, λ_0 equals 1.15 m, r equals 0.0055 m, Q_0 equals 5.0, $\Delta t/t$ equals 0.5 and α equals 10 dB/km. Based on these inputs, n equals 0.07 dispersed sardines per m^3. For the dispersed layer of sardines at sunrise on 5 September, f_0 equals 2.7 kHz, λ_0 equals 0.56 m, r equals 0.0041 m, Q_0 equals 7.0, and $\Delta t/t$ equals 0.8 and α equals 5.8 dB/km. Based on these inputs, n equals 0.05 dispersed sardines per m^3.

The absorption line at night may include contributions from schools of sardines ($f_0 \approx 0.9$ kHz at 20 m), dispersed anchovies ($f_0 \approx 1.1$ kHz at 9 m), and schools of anchovies, whereas the absorption line at sunrise on 5 September (at 2.7 kHz) appears to be due primarily to dispersed sardines. Significantly, this estimate is approximately equal to Barange *et al.*'s (1996) measurement of this parameter, viz., $0.12/m^3$.

The areal density, n_A, may be calculated from the expression:

$$n_A = nt. \tag{21}$$

Assuming that n equals $0.05/m^3$ and t equals 15 m, the areal density of dispersed sardines along the track is $0.75/m^2$. If the average weight of 16.5 cm long sardines is assumed to be 38 gm, then the biomass of dispersed adult sardines at this site at night is about 29 metric tons/km^2. This represents a lower bound on the biomass. It does not include the contribution due to sardines in schools. Calculation of the total biomass would require a model that considers the effect of close proximity between fish in schools on σ'_E.

For comparison, the echo sounder based estimate of the biomass of adult sardines in colonies in the Gulf of Lion is projected to be at least 32 metric tons per km^2 (disregarding errors associated with target strength and avoidance behavior).

These initial estimates of the biomass of adult sardines are comparable. The agreement, however, may to some extent be fortuitous. A more meaningful comparison will be possible when Ifremer completes its assessment of biomass from extensive echo sounder measurements at the Modal Lion site. Further research is needed to refine both methods, both estimates, and the uncertainties associated with these estimates.

XIV. SUMMARY AND CONCLUSIONS

This paper provides evidence of the effects of fish swim bladder resonance on transmission loss. Absorption losses exceeded 18 dB at a range of 12 km at the resonance frequency (1.3 kHz) of dispersed sardines at night. Absorption losses exceeded 35 dB at this range at the resonance frequency (~2.7 kHz) of dispersed sardines at sunrise. Measurements of the resonance frequencies of absorption lines were in good agreement with theoretical computations of the resonance frequencies of sardines, which were based on

nearly concurrently measured fish parameters. Matching measured and theoretical computations of the depth dependence of transmission loss permitted estimation of the depth of absorbing layers, which were in excellent agreement with echo sounder measurements of layer depths.

The estimated resonance frequencies of absorption lines associated with sardines in schools at a depth of about 65 m were substantially lower ($\sim$0.6) than the resonance frequencies of dispersed sardines at the same depth during daytime. This result is consistent with an approximate theoretical equation of the resonance frequency of the fundamental mode of a ''cloud'' of bubbles, an average separation between fish in schools of 0.8 fish lengths, and 5×10^3 sardines per school. These estimates are consistent with the previously reported geometrical properties of sardine schools, and the average separation between sardines in schools.

The resonance frequencies of other absorption lines are consistent with the hypothesis that absorptivity due to layers of sardines generally manifests itself in pairs of absorption lines, which correspond to the resonance frequencies of an ensemble of individual sardines, and the resonance frequencies of an ensemble of ''nuclei'' of schools.

The following topics require additional effort. The data should be analyzed at higher frequencies to permit investigation of the effects of juvenile sardines. A propagation model, which includes absorbing layers and mode coupling, should be developed. Distributions of the effective radius and eccentricity of swim bladders of sardines should be measured as a function of length. The magnitudes of f_0 and Q_0 of sardines should be measured as a function of season. A model of the extinction cross section of fish in the nucleus of schools should be developed. The analytical model of the resonance frequency of the school nucleus should be refined. The analytical procedure for estimating number density from absorptivity measurements should be refined. And possible effects of the high void content of schools of fish on sound speed and mode conversion should be investigated.

ACKNOWLEDGMENTS

Analysis of experimental results presented in this paper was supported by the Office of Naval Research. The experiment was supported by Nato's Saclant Undersea Research Centre. I would like to thank the engineering staff of NATO's Saclant Centre, in particular Reg Hollett, Sandro Barbagelatta, Enzo Michelozzi, Luigi Troiano and Alain Maguer, for their ingenuity and dedication during the experimental phase of this research program, Carlo Ferla for modifying the propagation code, Lothar Holtschmidt, Captain of the R.V. Alliance, for his cooperation, Pamela Nascetti and Andrea Perazzi for analyzing the echo sounder data, and Woody Nero of the Naval Research Laboratory (NRL) for theoretical computations. I would also like to thank David Bradley, former Director of NATO's Saclant Centre, Regis Martin Lauzer, Director of GESMA, France, Eric Hartwig, Director of the Ocean and Atmospheric Science and Technology Directorate (NRL), Ed Franchi, Superintendent of the Acoustics Division (NRL) and Marshall Orr, Head of the Signal Processing Branch (NRL), for their support and encouragement. In addition, I would like to thank Bernard Liorzou and Carla Scalabrin of IFREMER, France, and Massimo Azzali of IRPEM, Italy for their technical support, and David Weston and Van Holliday for many stimulating discussions and useful insights. I am grateful to Rick Love, Tom Hayward, Peter Mignerey, Steve Wales and Steve Wolf of the Naval Research Laboratory, and to Paul Smith of the National Marine Fisheries Service, for many helpful discussions, and their reviews of this paper. I would also like to thank the formal reviewers for their thoughtful comments and helpful suggestions.

Andreeva, I. (**1964**). ''Scattering of sound by air bladders of fish in deep sound scattering layers,'' Sov. Phys. Acoust. **10**, 17–20.

Anonymous. (**1981**). ''Atlas of the living resources of the seas,'' FAO Fisheries Department, Food and Agriculture Organization of the United Nations, Rome.

Aoki, I., and Inagaki, T. (**1988**). ''Photographic observations on the behaviour of Japanese anchovy Engraulis japonica at night in the sea,'' Marine Biology Progress Series **43**, 213–221.

Azzali, M., *et al.* (**1985**). ''Relationship between the forms of pelagic fish distribution and nycthermal periods: a tentative model,'' Oebalia **11**, 471–488.

Azzali, M., *et al.* (**1995**). ''Valutazione delle biomasse, delle loro composizioni specifiche e distribuzioni spaziali, utilizzando una metodologia acoustica integrata,'' C.N.R. I.R.Pe.M., Ancona, Italy.

Baltzer, W., and Pickwell, G. (**1970**). ''Resonant acoustic scattering from gas-bladder fishes,'' in Proceedings of an International Symposium on Biological Scattering in the Ocean, edited by B. Farquhar, Maury Center for Ocean Science, Washington, D.C.

Barange, M., *et al.* (**1996**). ''Empirical determination of in situ target strengths of three loosely aggregated pelagic fish species,'' ICES J. Mar. Science **53**, 225–231.

Blaxter, J., and Batty, R. (**1990**). ''Swimbladder behaviour and target strength,'' Rapp. P.-V. Reun.-Cons. Int. Explor. Mer. **189**, 233–244.

Blaxter, J., and Hunter, J. (**1982**). ''The biology of clupeoid fishes,'' in *Advances in Marine Biology* (Academic, New York), Vol. 20, pp. 1–223.

Brawn, V. (**1969**). ''Buoyancy of Atlantic and Pacific herring,'' J. Fish. Res. Board Can. **26**, 2077–2091.

Brekhovskich, L. (**1960**). *Waves in Layered Media* (Academic, New York).

Carey, W., and Roy, R. (**1993**). ''Sound scattering from microbubble distributions near the surface,'' in *Ocean Reverberation*, edited by D. Ellis, J. Preston, and H. Urban (Kluwer, Dordrecht).

Ching, P., and Weston, D. (**1971**). ''Wide band studies of shallow water acoustic attenuation due to fish,'' J. Sound Vib. **18**, 499–510.

Ching, P., and Weston, D. (**1975**). ''Fast acoustic fluctuations caused by fish,'' J. Sound Vib. **39**, 287–292.

Clay, C. (**1991**). ''Low resolution acoustic scattering models: Fluid filled cylinders and fish with swim bladders,'' J. Acoust. Soc. Am. **89**, 2168–2179.

Clay, C., and Medwin, H. (**1977**). *Acoustical Oceanography: Principles and Applications* (Wiley, New York).

Cullen, J., Shaw, E., and Baldwin, H. (**1963**). ''Methods for measuring the three dimensional structure of fish schools,'' Anim. Behav. **13**, 534–543.

Culley, M. (**1971**). *The Pilchard: Biology and Exploitation* (Pergamon, New York).

d'Agostino, L., and Brennan, C. (**1988**). ''Acoustical absorption and scattering cross sections of spherical bubble clouds,'' J. Acoust. Soc. Am. **84**, 2126–2134.

Diachok, O. (**1996**). ''Fish absorption spectroscopy,'' in Proceedings of the 3rd European Conference on Underwater Acoustics, edited by J. Papadakis (E.C. Press, Luxembourg).

Diachok, O., and Ferla, C. (**1996**). ''Measurement and simulation of the effects of absorptivity due to fish on transmission loss in shallow water,'' in Proceedings of the Meeting of the Ocean Engineering Society, pp. 524–529 (IEEE Press, Bellingham, WA).

Druzhinin, O., *et al.* (**1996**). ''Low frequency acoustic wave generation in a resonant bubble layer,'' J. Acoust. Soc. Am. **100**, 3570–3580.

Diner, N., Personal Communication, IFREMER, France, 1996.

Ferla, C., Porter, M., and Jensen, F. (**1993**). Report SM-274, Saclant Undersea Research Centre, La Spezia, Italy.

Feuillade, C. (**1995**). ''Scattering from collective modes of air bubbles in water and the physical mechanism of superresonances,'' J. Acoust. Soc. Am. **98**, 1178–1190.
Feuillade, C., and Werby, M. (**1994**). ''Resonances of deformed gas bubbles in liquids,'' J. Acoust. Soc. Am. **96**, 3684–3692.
Feuillade, C., *et al.* (**1996**). ''A low frequency acoustic scattering model for small schools of fish,'' J. Acoust. Soc. Am. **99**, 196–208.
Foote, K. (**1987**). ''Fish target strength for use in echo integrator surveys,'' J. Acoust. Soc. Am. **82**, 981–987.
Freon, P., *et al.* (**1993a**). ''Consequences of fish behaviour for stock assessment,'' ICES Marine Science Symposium, , 190–195.
Freon, P., *et al.* (**1993b**). ''Diurnal variation in fish density,'' Aquatic Living Resources **6**, 220–234.
Freon, J., *et al.* (**1996**). ''Diel variability of school structure with special reference to transition periods,'' ICES J. Mar. Science, **53**, 459–464.
Guennegan, Y., *et al.* (**1997**). ''Suivi de l'exploitation et de la ressource des petits pelagiques du Golfe du Lion,'' DRV-97-RH, Ifremer, Sete, France.
Graves, J. (**1977**). ''Photographic method for measuring the spacing and density within pelagic fish schools at sea,'' Fishery Bulletin, U.S. **75**, 230–234.
Herrera, J., and Munoz, F. (**1957**). ''Consideraciones biologicas sobre composicion quimica de la sardina (Sardina pilchardus Walb.) de Castallon, Inv. Pesq.,'' **VII**, 33–43.
Hersey, J., *et al.* (**1962**). ''Sound scattering spectra of deep scattering layers in the western north Atlantic Ocean,'' Deep-Sea Res. **8**, 196–210.
Hickling, C. (**1945**). ''The seasonal cycle in the Cornish Pilchard,'' J. Mar. Biol. Assoc. U.K. **26**, 115–138.
Holliday, D. (**1972**). ''Resonance structure in echoes from schooled pelagic fish,'' J. Acoust. Soc. Am. **51**, 1322–1332.
Holliday, V. (**1978**). ''Data analysis and summary, MORDAX II,'' Tracor Document Number T-78-SD-002-U, Tracor, Inc.
Hunter, J., NMFS, personal communication, 1997.
Kustov, L., *et al.* (**1982**). ''Parametric source radiation in a bubble layer,'' Acoust. Lett. **6**, 15–17.
Liorzou, B., IFREMER, France, personal communication, 1995.
Love, R. (**1970**). *The Chemical Biology of Fishes* (Academic, New York).
Love, R. (**1978**). ''Resonant scattering by swimbladder bearing fish,'' J. Acoust. Soc. Am. **64**, 571–580.
Loukashkin, A. (**1970**). ''On the diet and feeding behavior of the northern anchovy,'' Engraulis mordax, Proc. Cal. Acad. Sci. **37**, 419–458.
Lovik, A., and Hovem, J. (**1979**). ''An experimental investigation of swimbladder resonances in fishes,'' J. Acoust. Soc. Am. **66**, 850–854.
Lu, N., *et al.* (**1990**). ''Underwater noise emissions from bubble clouds,'' IEEE J. Ocean Eng. **15**, 275–285.
Millot, C., and Crepon, M. (**1981**). ''Inertial oscillations on the continental shelf of the Gulf of Lions—observations and theory,'' J. Phys. Oceanogr. **11**, 639–657.
Minnaert, F. (**1933**). ''On musical air bubbles and the sounds of running water,'' Philos. Mag. **16**, 235–248.
Misund, O. (**1991**). ''Dynamics of moving masses: Variability in packing density, shape and size among pelagic schools,'' in *Swimming Behaviour of Schools Related to Fish Capture and Acoustic Abundance Estimation*, Ph.D. Thesis, Bergen, Norway.
Misund, O., and Algen, A. (**1992**). ''Swimming behaviour of fish schools in the North Sea during acoustic surveying and pelagic trawl sampling,'' ICES J. Mar. Sci. **49**, 325–334.
Misund, O., *et al.* (**1995**). ''Mapping the shape, size and density of fish schools by echo integration and a high resolution sonar,'' ICES J. Marine Science **52**, 11–20.
Moffet, M., and Mellen, R. (**1977**). ''Model for parametric acoustic sources,'' J. Acoust. Soc. Am. **61**, 325–337.
Mullin, M. (**1993**). ''Webs and Scales: Physical and Ecological Processes in Marine Recruitment,'' University of Washington Press, Seattle.
Nero, R. (**1996**). ''Model estimates of acoustic scattering from schools of large yellowfin tuna,'' Naval Research Laboratory Report, NRL/MR/174-95-7708., Stennis Space Center, MI.
Nero, R., NRL, personal communication, 1996.
Nonacs, P., *et al.* (**1988**). ''Modeling foraging in the northern anchovy (Engraulis mordax): individual behavior can predict school dynamics and population biology,'' Can. J. Fisheries and Aquatic Sciences, **55**, 1179–1188.
Ona, E. (**1990**). ''Physiological factors causing natural variations in acoustic target strengths of fish,'' J. Mar. Biol. Assoc. U.K. **70**, 107–127.
Pichot, P., and Aldebert, Y. (**1978**). ''La Peche de la sardine en Mediterranee Francaise,'' Science et Peche, Bull. Inst. Peches marit., No. 277, , 5.
Pitcher, T., and Parrish, J. (**1993**). ''Functions of shoaling behaviour in teleosts,'' in *Behaviour of Teleost Fishes*, edited by T. Pitcher (Chapman and Hall, London).
Pitcher, T., *et al.* (**1985**). ''Schooling mackerel and herring choose neighbors of similar size,'' Marine Biology **86**, 319–322.
Popper, A. (**1997**). ''A clupeid fish can detect ultrasound,'' Nature **389**, 341.
Price, J., *et al.* (**1986**). ''Diurnal cycling: Observations and models of the upper ocean response to diurnal heating, cooling, and wind mixing,'' J. Geophys. Res. **91**, 8411–8427.
Radakov, D. (**1973**). *Schooling in the Ecology of Fish* (Wiley, New York).
Rose, G. (**1992**). ''A review of problems and new directions in the applications of fisheries acoustics on the Canadian East Coast,'' Fisheries Research, **14**, 105–128.
Rouseff, D., and Ewart, T. (**1995**). ''Effect of random sea surface and bottom roughness on propagation in shallow water,'' J. Acoust. Soc. Am. **98**, 3397–3404.
Scalabrin, C., IFREMER, France, personal communication, 1996.
Scalabrin, C. (**1997**). ''Identification acoustique des especes pelagiques du sujet suivant,'' Ph.D. Thesis, Brest, France.
Scalabrin, C., *et al.* (**1996**). ''Narrow band acoustic identification of monospecific fish shoals,'' ICES Journal of Marine Science **53**, 181–188.
Schindall, J. (**1995**). ''Acoustic scattering from compact bubble clouds,'' Ph.D. Thesis, University of Mississippi.
Schneider, H. (**1990**). ''Average sound intensities in randomly varying sound speed structures,'' in *Ocean Variability and Acoustic Propagation*, edited by J. Potter and A. Warn Varnas (Kluwer, Dordrecht).
Sellschopp, J. (**1990**). ''Stochastic ray tracing in thermoclines,'' in *Ocean Variability and Acoustic Propagation*, edited by J. Potter and A. Warn Varnas (Kluwer, Dordrecht).
Smith, P., personal communication, 1997.
Smith, P. (**1978**). ''Precision of sonar mapping for pelagic fish assessment in the California current,'' J. Cons. Explor. Mer. **38**, 33–40.
Smith, P., *et al.* (**1989**). ''Analysis of patterns of distributions of zooplankton aggregations from an acoustic doppler current profiler,'' CalCOFI Report, p. 30.
Soria, M., *et al.* (**1996**). ''Analysis of vessel influence on spatial behaviour of fish schools using multi-beam sonar and consequences for biomass estimates by echo sounder,'' ICES J. Mar. Science **53**, 453–458.
Sundness, G., and Sand, O. (**1975**). ''Studies of a physotome swimbladder by resonance frequency analysis,'' J. Cons. Int. Eplor. Mer. **36**, 176–182.
Thiele, R. (**1990**). ''Modeling of sound propagation in a randomly varying ocean by stochastic mode coupling,'' in *Ocean Variability and Acoustic Propagation*, edited by J. Potter and A. Warn Varnas (Kluwer, Dordrecht).
Wardle, C. (**1983**). ''Fish reactions to towed gears,'' in *Experimental Marine Biology at Sea*, edited by A. MacDonald and I. Priede (Academic, London), pp. 167–195.
Weston, D. (**1967**). ''Sound propagation in the presence of bladder fish,'' in *Underwater Acoustics*, Vol. 2, edited by V. Albers (Plenum, New York), pp. 55–58.
Weston, D. (**1970**). ''Sound extinction by fish in one-way shallow water propagation'' in *Proceedings of an International Symposium on Biological Sound Scattering in the Ocean*, edited by B. Farquhar (Maury Center for Ocean Science, Washington, D.C.).
Whitney, R. (**1969**). ''Schooling of fishes relative to available light,'' Trans. Am. Fisheries Society **98**, 497–504.
Woodhead, P. (**1966**). ''The behavior of fish in relation to light in the sea,'' Oceanogr. Mar. Biol. Ann. Rev. **4**, 337–403.
Ye, Z. (**1996**). ''On acoustic attenuation by swimbladder fish,'' J. Acoust. Soc. Am. **100**, 670–672.
Ye, Z., and Farmer, D. (**1994**). ''Acoustic scattering from swimbladder fish at low frequencies,'' J. Acoust. Soc. Am. **96**, 951–956.
Zhou, J., *et al.* (**1991**). ''Resonant interaction of sound wave with internal solitons in the coastal zone,'' J. Acoust. Soc. Am. **90**, 2042–2054.
Zupanovich, S. (**1967**). ''Study of sardine (sardina pilchardus) behavior in their natural environment by echo sounder and environmental factors,'' FAO Conference on fish behavior, Fish and Agriculture Organization of the United Nations, Report 62.2, pp. 269–282.

Azimuthal dependence of Bragg scattering from the ocean surface

C. Scott Hayek, Iman W. Schurman, John H. Sweeney, and C. Allan Boyles
Johns Hopkins University, Applied Physics Laboratory, Johns Hopkins Road, Laurel, Maryland 20723

(Received 12 March 1997; accepted for publication 16 November 1998)

Frequency selective scattering of water-borne acoustic waves by the rough sea surface (Bragg scattering) has been observed, in particular during the Critical Sea Test series. The directional nature of the gravity wave spectrum observed by Mitsiyasu, Donelan, Banner and others implies that the interface scattering will be directional, that is, depend upon azimuth relative to the receiver. During the ocean exercise Critical Sea Test 4, Bragg scattering was observed at 250 Hz over a wide range of azimuths using a linear hydrophone array with approximately one degree azimuthal resolution. The amplitude of the Bragg scattering and its dependence on azimuth closely matched model predictions based on the Donelan 2-D wave spectrum, first-order perturbation theory and a normal mode reverberation model. © *1999 Acoustical Society of America.* [S0001-4966(99)05503-4]

PACS numbers: 43.30.Gv, 43.30.Hw [DLB]

INTRODUCTION

In this paper we report results derived from probing the air–sea interface with a narrow-band acoustic signal. The advantage of the narrow-band technique is that the air–sea interface can be observed directly by measuring the upward and downward Bragg-shifted echo. The azimuthal dependence of the intensity of the shifted scattered sound is a direct consequence of the two-dimensional gravity wave field. The study's goal was to investigate the consistency between the scattered sound levels and an established two-dimensional gravity wave model.

The experiment described in this paper was performed in April 1990 as part of the Critical Sea Test Program, in coordination with the surface scattering experiments devised and conducted by F. T. Erskine and P. M. Ogden.[1,2] Their technique used explosive sources to investigate frequencies from 70 Hz to 1500 Hz. These short range, broadband underwater acoustic experiments, which show the transition from air–sea boundary scattering to volumetric scattering as a function of frequency and wind speed, have been described in Refs. 3–5.

The link between the two-dimensional gravity wave field and acoustic reverberation comes about because backscattering of sound from the ocean surface occurs predominantly at the Bragg frequencies. Surface reverberation arises due to the selective reinforcement of scattering from the gravity wave matched to the half-wavelength of the incident acoustic wave, just as an optical diffraction grating reinforces a single wavelength in a given viewing geometry. This ''matched'' gravity wave is part of the continuum of wind waves traveling predominantly in the direction of the wind. Nonlinear interactions between gravity waves create waves traveling in directions other than with the wind. The energy of these off-the-wind directional waves differs as azimuthal angle varies. Two-dimensional water wave spectra models have been published to quantify the total wave field's directional dependence.[6,7]

Perturbation theory connects the directional dependence of gravity wave spectra with azimuthal dependence of surface scattering strength. At a given frequency the scattering strength is proportional to the gravity wave spectrum evaluated at the Bragg wavelength. Thus, the azimuthal dependence of the two-dimensional gravity wave spectrum is imparted to the acoustic scattering coefficient.

The narrow-band surface reverberation experiment reported on here isolated the interface-scattered acoustic signal using Doppler shift due to the motion of the scattering gravity waves. For this investigation 250 Hz gated sinusoids of sufficient duration to resolve the up and down shifted Bragg reflection from the volume reflections near zero Doppler were transmitted. Once the up and down Bragg scattered energy had been isolated and plotted as a function of azimuth (receiver beam angle) the surface scattered intensities could be compared to the result predicted from the two-dimensional gravity wave spectrum and perturbation theory.

Section I presents the models used for the Bragg reverberation predictions, including the perturbation theory used to derive the scattering strength and the normal mode reverberation model. Section II describes the data measurement and analysis techniques. Section III compares the measured Bragg data to the modeled results.

I. MODEL DESCRIPTION

A. First-order perturbation theory scattering strength

Recent work has been done to calculate the scattered field from a rough sea surface using perturbation theory. In Refs. 8 and 9 a static sea surface with a symmetric sea surface spectrum was used. While a moving sea surface was used in Ref. 10, the sea surface spectrum was also assumed to be symmetric, which means that the energy in the gravity wave of wavelength Λ traveling in azimuthal direction ϕ_D is equal to the energy of a wave of the same wavelength traveling in the opposite direction. This assumption would mean that a surface wave component traveling against the wind has an amplitude equal to the component traveling with the wind—clearly refuted by observation. The approaches in

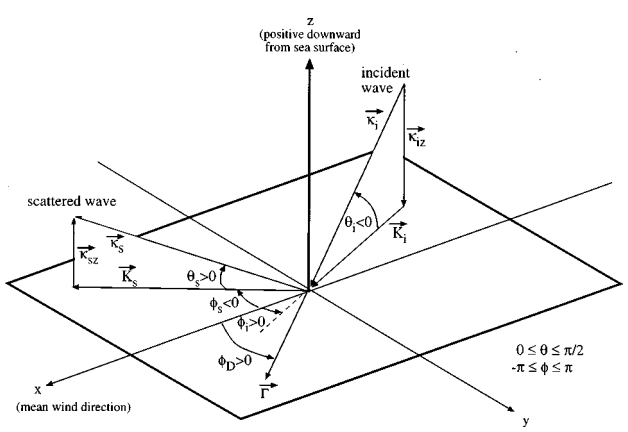


FIG. 1. Scattering geometry and angle conventions used in derivation of scattering strength.

Refs. 8–10, therefore, do not account for the asymmetry of the 2-D sea surface spectrum, which this experiment was designed to measure.

The scattering cross section for an asymmetric, moving sea surface has been known in the radar community for some time.[11] Most references, such as Ref. 11, assume an infinite rough surface and an incident plane wave, but give no detailed derivation. Since an infinite rough surface and an incident plane wave are physically unrealizable, and it is mathematically impossible to derive a scattering strength for this case, we present in Appendix A a derivation of the scattering cross section for an asymmetric, moving rough surface for a practical situation. Essentially, it amounts to illuminating a finite area of the surface with a beam. The derivation in Appendix A is an extension of the work of E. I. Thorsos to three-dimensional scattering. It shows in detail the approximations that must be made to go from the ideal (but unrealizable) case of an infinite rough surface with an incident plane wave, to scattering from a finite surface area by a nonideal plane wave.

Figure 1 shows the scattering geometry. The *z*-axis is taken to be in the direction of increasing water depth. The figure is inverted for clarity. All angles labeled θ with any subscript are vertical grazing angles measured from the *xy*-plane. The angle θ_i for the incident wave is negative, while the scattering angle θ_s is positive. All angles labeled ϕ with any subscript are azimuthal angles measured in the *xy*-plane. The positive *x*-axis is taken to be $\phi=0$, with ϕ being positive with range $0 \leqslant \phi \leqslant \pi$ if the direction of rotation is from the positive *x*-axis toward the positive *y*-axis, and ϕ is negative with range $-\pi \leqslant \phi \leqslant 0$ if the direction of rotation is from the positive *x*-axis toward the negative *y*-axis. The mean wind direction is taken to be in the positive *x*-direction. The angle ϕ_i is the azimuthal angle for the incident plane wave.

The scattering coefficient σ is given by Eq. (A67) in Appendix A as

$$\sigma = 2\kappa_{iz}^2 \kappa_{sz}^2 \{ W(\mathbf{K}_s - \mathbf{K}_i)\, \delta(\omega_s - \omega_i - \Omega_\Gamma) + W(-\mathbf{K}_s + \mathbf{K}_i)\, \delta(\omega_s - \omega_i + \Omega_\Gamma) \}, \tag{1}$$

where ω_s = frequency of scattered wave, ω_i = frequency of incident wave, $\mathbf{K}_s$ = horizontal component of the scattered wave vector, $\mathbf{K}_i$ = horizontal component of the incident wave vector, κ_{iz} = vertical component of the incident wave vector, κ_{sz} = vertical component of the scattered wave vector, W = directional wave number spectrum, and δ = Dirac delta function. The dispersion relation for the gravity waves is taken to be $\Omega_\Gamma = \sqrt{g|\mathbf{\Gamma}|}$, where g is the acceleration due to gravity and $\mathbf{\Gamma}$ is the gravity wave number vector.

B. Donelan water wave spectra

The Donelan frequency-direction spectrum[6] is given by

$$F(\Omega, \phi_D) = S(\Omega) h(\phi_D; \Omega), \quad -\pi \leqslant \phi_D \leqslant \pi, \tag{2}$$

where $S(\Omega)$ is the frequency spectrum, and the spreading function h determines the directional distribution of the surface wave energy. Here ϕ_D measures the azimuthal angle relative to the mean wind direction, and Ω is the angular frequency of the gravity waves. Letting Ω_p denote the peak spectral frequency, c_p the wave speed at frequency Ω_p, and U_c the wind speed at a height of 10 m, the Donelan frequency spectrum is

$$S(\Omega) = \alpha g^2 \Omega^{-5} (\Omega/\Omega_p) \exp[-(\Omega_p/\Omega)^4] \gamma^\epsilon, \tag{3}$$

where

$$\alpha = 0.006 (U_c/c_p)^{0.55}, \quad 0.83 < U_c/c_p < 5, \tag{4}$$

$$\gamma = 1.7, \quad \text{for } 0.83 < U_c/c_p < 1,$$

$$= 1.7 + 6.0 \log_{10}(U_c/c_p), \quad 1 \leqslant U_c/c_p < 5, \tag{5}$$

$$\epsilon = \exp[-(\Omega - \Omega_p)^2/(2\nu^2\Omega_p^2)], \tag{6}$$

and

$$\nu = 0.08[1 + 4/(U_c/c_p)^3] \quad 0.83 < U_c/c_p < 5. \tag{7}$$

The spreading function h is given by

$$h(\phi_D; \Omega) = \frac{\beta}{2} \operatorname{sech}^2(\beta \phi_D), \quad -\pi \leqslant \phi_D \leqslant \pi, \tag{8}$$

where β is a dimensionless parameter that controls the anisotropy of surface wave amplitudes. In general β is a function of the dimensionless frequency Ω/Ω_p and the wave age c_p/U_c.

Since as a wave ''ages'' it grows longer and as it grows longer it moves faster, the wave age is greater for an older wave and is thus a dimensionless measure of the wave age. Note that wind driven water waves cannot travel faster than the wind speed.

The spreading function h is normalized by the condition

$$\int h(\phi_D; \Omega)\, d\phi_D = 1. \tag{9}$$

Let $W(\mathbf{\Gamma}, \Omega)$ denote the complete directional wavenumber spectrum which is defined by Eq. (A38) in Appendix A. The frequency directional spectrum F is related to the wavenumber directional spectrum W by the relation[6,7]

$$F(\Omega, \phi_D) = 2 \int_0^\infty W(\Gamma, \phi_D, \Omega) \Gamma\, d\Gamma. \tag{10}$$

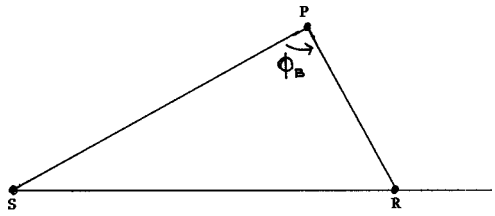


FIG. 2. Geometry for bistatic scattering.

If we now substitute W into Eq. (10), change the variable of integration to $u=\sqrt{g\Gamma}$, and integrate, to get

$$F(\Omega,\phi_D)=\frac{2}{g}\int_0^\infty [W(u^2/g,\phi_D,\Omega)\,\delta(\Omega-u) + W(u^2/g,\phi_D+\pi,\Omega)\,\delta(\Omega+u)]u^3\,du = \frac{2\Omega^3}{g^2}W(\Omega^2/g,\phi_D). \tag{11}$$

From Eq. (11) we have

$$W(\mathbf{\Gamma},\Omega)=W(\Gamma,\phi_D,\Omega)=\frac{g^{1/2}}{2\Gamma^{3/2}}F(\sqrt{g\Gamma},\phi_D). \tag{12}$$

It is shown by Warfield[12] that for three-dimensional scattering the first-order Bragg frequency shift ω_B associated with incident and scattered grazing angles θ_i and θ_s can be written in the form

$$\omega_B=[kg(\cos^2\theta_s+\cos^2\theta_i+2\cos\theta_s\cos\theta_i\cos\phi_B)^{1/2}]^{1/2}, \tag{13}$$

where ϕ_B is the bistatic angle shown in Fig. 2, and $k=\kappa_s=\kappa_i$. In Fig. 2 which is a projection onto the xy-plane, S is the source, R is the receiver and P is the scattering point. Since our source and receiver are co-located, the bistatic angle $\phi_B=0$. Equation (13) then becomes

$$\omega_D(\theta_i,\theta_s)=|\omega_s-\omega_i|=g^{1/2}(\kappa_i\cos\theta_i+\kappa_s\cos\theta_s)^{1/2}. \tag{14}$$

In the remainder of the paper we will drop the subscript Γ on the gravity wave frequency Ω and understand that it is the frequency constrained by the dispersion relation Eq. (A43).

Using Eq. (14), Eq. (2), and Eq. (12), Eq. (1) can be put into the following forms for the scattering strength amplitudes at the down- and upshifted Bragg frequencies. Defining

$$\sigma_v(\omega_D)=\frac{\kappa_{iz}^2\kappa_{sz}^2}{\omega_D^3}g^2S(\omega_D), \tag{15}$$

we have for the downshifted scattering strength

$$\sigma^{\text{down}}(\omega_D,\phi_i)=\sigma_v(\omega_D)h(\phi_i;\omega_D). \tag{16}$$

We have replaced ϕ_D in the argument of the function h by ϕ_i since for the downshifted case ϕ_i will always be in the same direction as ϕ_D for the entire range $-\pi\leqslant\phi_D\leqslant\pi$.

For the upshifted scattering strength two separate cases must be considered. However, for both cases, the incident wave direction ϕ_i is in the opposite direction of the gravity waves ϕ_D.

Thus when $-\pi\leqslant\phi_D\leqslant 0$, we have $\phi_D=\phi_i-\pi$ so we can write

$$\sigma^{\text{up}}=\sigma_v(\omega_D)h(\phi_i-\pi;\omega_D). \tag{17}$$

When $0\leqslant\phi_D\leqslant\pi$, we have $\phi_D=\phi_i+\pi$, so we can write

$$\sigma^{\text{up}}=\sigma_v(\omega_D)h(\phi_i+\pi;\omega_D). \tag{18}$$

C. Normal mode reverberation model

A normal mode approach is used to model the observed Bragg reverberation. A normal mode model was used instead of a ray model for the following reason: the strong convergence zone environment. A convergence zone possesses caustics and shadow zone diffraction which are not handled well by a ray model. The model we used is based upon an unpublished model developed by K. J. McCann at JHU/APL in 1989.

The reverberation equation used for the Bragg reverberation predictions is

$$\text{RL}(r,\phi)=\text{SL}+10\log_{10}[I(r,\phi)/I_0] - 10\log_{10}[180\cdot w(\phi)/\pi], \tag{19}$$

where $\text{RL}(r,\phi)$ is the reverberation level ($\text{dB}//\mu\text{Pa}^2$) received in an equivalent 1° beam, SL is the source level ($\text{dB}//\mu\text{Pa}^2$ @1 m), $I(r,\phi)$ is the reverberation intensity due to a surface scattering patch at (r,ϕ), I_0 is the equivalent plane wave intensity of the transmitted cw pulse at a reference distance of 1 m, and $w(\phi)$ is the receive array effective beamwidth (in radians) at steering ϕ.

The beam reverberation intensity I is modeled as

$$I(r,\phi)=I_0\int\int_{A(r)}B(\phi,\phi')|P_{\text{rec}}(r',\phi')|^2\,dA, \tag{20}$$

where $B(\phi,\phi')$ is the normalized beampattern of the receive array, $A(r)$ is the scattering patch at the range r, and $P_{\text{rec}}(r,\phi)$ is the relative pressure of the reverberation received at the acoustic center of the horizontal receive array. Equation (20) expresses the beam intensity as an incoherent sum of contributions from all azimuths (assumed to be statistically independent). The received pressure P_{rec} includes the effects of propagation from the steered source array to the scattering patch, scattering at the surface, and propagation back to the receiver. The received intensity $|P_{\text{rec}}|^2$ thus accounts for a round trip transmission loss. The scattering patch $A(r)$ is approximated by an annulus with range extent $c\tau/2$, where c is an effective sound speed and τ is the pulse duration.

The range component of the received pressure is given by

$$P_{\text{rec}}(r)=\sum_n\sum_m P_n^{\text{inc}}(r,z_s)\sqrt{\sigma_v(\Omega_{nm})}P_m^{\text{scatt}}(r,z_r), \tag{21}$$

with

$$\Omega_{nm}=(gk_i)^{1/2}(\cos\theta_n+\cos\theta_m)^{1/2}. \tag{22}$$

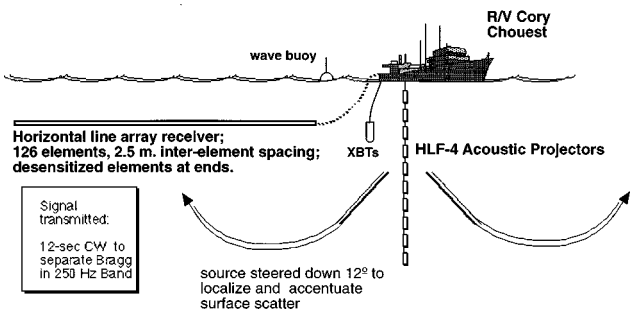


FIG. 3. The R/V CORY CHOUEST was used to tow both the vertical array of 10 HLF-4 projectors and the horizontal line array receiver. Supporting wind, wave and water sound speed profile measurements were taken during the Bragg scattering experiment.

Here P_n^{inc} is the pressure of the wave incident on the surface, and P_m^{scatt} is the pressure of the wave scattered by the surface.

A derivation of Eq. (21) is presented in Appendix B.

II. THE EXPERIMENT

A. Data collection

The experiment described here was part of Critical Sea Test 4 conducted in the Gulf of Alaska in April 1990. The research ship R/V CORY CHOUEST deployed the vertical source array (VLA) consisting of 10 HLF-4 projectors spaced 3.66 m apart with an acoustic center 163 m below the surface. For receiving, a 126 element (2.5 m inter-element spacing) horizontal line receiving array (HLA) was deployed and towed 137 m below the surface (see Fig. 3). Tow speed was approximately 1.7 m/s. The separation between VLA and the HLA was 2.3 km. One forward and one aft hydrophone of the HLA were desensitized to allow measurement of the signal transmitted from the VLA. Wind speed and wave height measurements were taken throughout the exercise by ship mounted anemometers and an ENDECO Type 1156 wavebuoy, respectively. Sound speed profiles were obtained during the data collection using XBT and CTD devices.

During the experiment 12 s long, nominal 250 Hz tones with Hanning time weighting were transmitted, allowing clear resolution of the 0.7 Hz Bragg shift associated with interface scattering. The time interval between pings was 900 s. Reverberation was received on the HLA which resolved angular sectors of about 1° at broadside and about 10° at endfire.

During the test the sound speed increased monotonically from just below the sea surface to the bottom (Fig. 4). Water depth was approximately 5000 m and wind speed was 6.8 m/s, arriving from 310° true. The heading for the CORY CHOUEST was 83° true.

The source vertical aperture helped maximize the contrast between the surface scattered signal and background. At the 250 Hz transmit frequency the source array had a beamwidth of approximately 9°. Steering the source beam down 12° produced a distinct reverberation ''hump'' from the first convergence zone or ''pseudo convergence zone'' (Fig. 5). Note that the bottom and surface are strongly ensonified at separate ranges (times): in particular, when the surface and

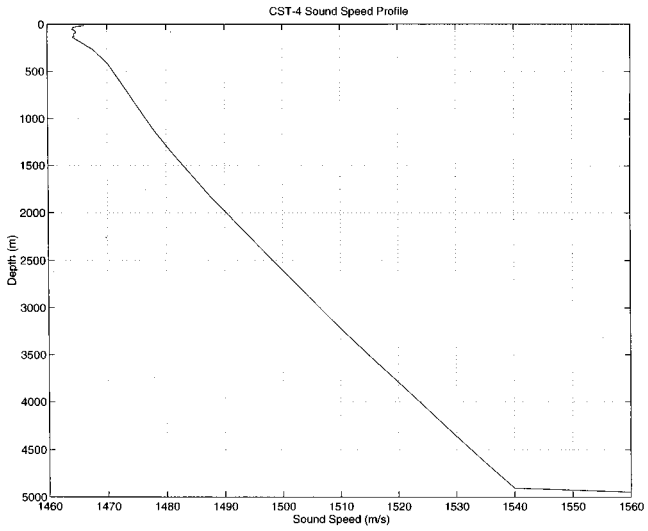


FIG. 4. Representative sound speed profile measured during the Bragg reverberation experiment.

near-surface are ensonified the bottom illumination is highly attenuated. Thus, the surface reverberation should have a high signal to background ratio.

Steering down 12° also served to increase the surface scattering strength relative to a 0° source steering angle. The surface incident angle for 12° steering was predicted to be 12.5°, while the incident angle for 0° steering would have been 3°. At 6 m/s wind speed, this corresponds to an approximate 25 dB scattering strength advantage.

B. Data processing and analysis

Data from the desensitized hydrophone were FFT'd to determine the source spectrum and level. The level was determined to be 226.1 d dB//μPa^2.

Calibrated hydrophone data from the 126 hydrophone towed array were beamformed. This yielded 126 beams (approximately 1° sector width at broadside) covering forward endfire to aft endfire. The HLA had seven inoperative hydrophones during the data collection. Time series from four transmissions were compensated for own-ship doppler, and then spectra covering the return time interval were extracted.

Achieving consistency between measured quantities and model calculations was of paramount importance. The data were in the form of spectral levels in 1/16 Hz bins, each representing a single 12 s FFT.

To compare experiment and model, spatial, time/range, frequency, and noise effects had to be accounted for in the data. The required data ''transforms'' are discussed below.

In the spatial transform, the data were scaled to a constant equivalent beamwidth of 1°.

In the time/range transform the data were incoherently averaged over two overlapped FFT transforms, resulting in an effective time-width of 12 s.

The frequency transform gathered data from several spectral bins to represent the energy in the return. The algorithm created 1/12 Hz levels oversampled such that they are on 1/16 Hz centers. The required correction for overcounting energy when adding bins was applied. Due to shading of the

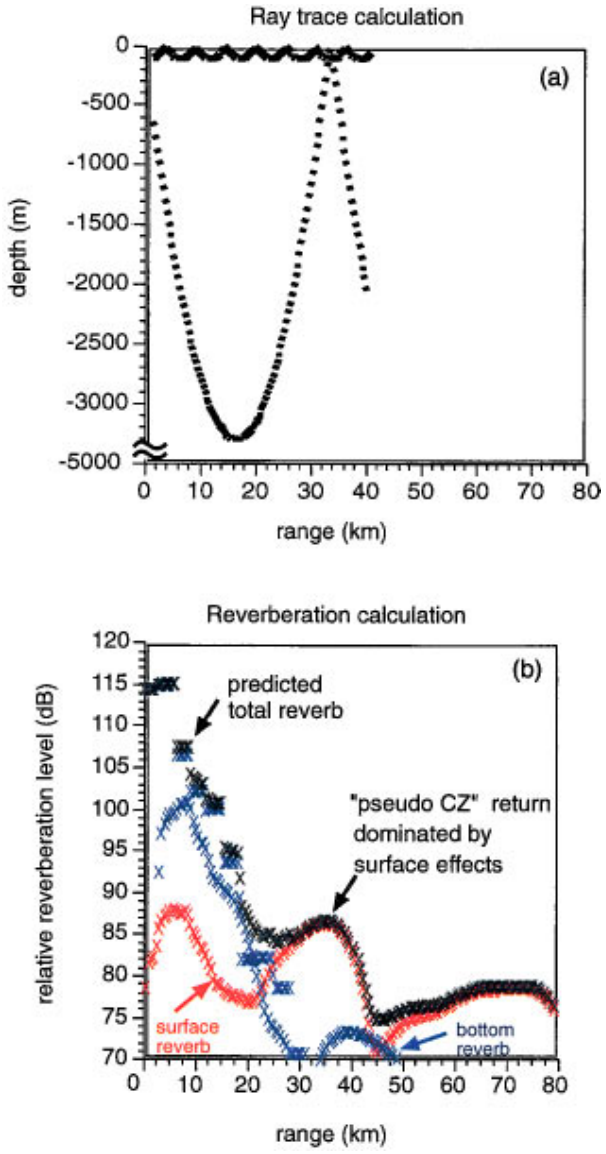


FIG. 5. Vertically steering the source beam 12 degrees down in the half-channel sound speed profile accentuates the surface return by steepening the incident grazing angle at the ''pseudo-CZ'' range, while still avoiding detrimental bottom returns.

transmitted waveform and small errors in the assumed instantaneous ship speed, the Bragg returns in the data occupied more than a single 1/16 Hz bin. To gather the entire Bragg return, a sum over five 1/16 Hz bins was computed.

Data from four of the transmissions were averaged.

The noise transform was performed to reduce the effect of noise on the estimates of Bragg signal energy. To estimate the impact of beam noise on the Bragg return measurements, a beam-by-beam estimate of the signal plus noise to noise ratio was prepared. This was done by comparing the averaged Bragg estimates to two estimates of the underlying beam noise at the Bragg frequencies. One estimate was derived from nearby, out-of-band noise data, while the other was derived from an estimate of the side lobe contamination at the Bragg frequencies due to the reverberation returns at those same frequencies on other azimuths.

On any given beam the larger of the two noise estimates was used together with the Bragg signal plus noise estimate to create a signal only measurement. For convenience, these were referred back to the original signal plus noise estimate to create a noise estimate in dB. These corrections varied widely, but mainly fell into the 1–2.5 dB range. By happenstance, the strength of the own ship Doppler sidelobes was strong in the strong signal directions, and limited the Bragg signal-to-noise ratio to about 6 dB. The weak Bragg returns still have a 2–3 dB signal-to-noise ratio above the ambient beam noise in those regions where the own ship Doppler sidelobes are absent.

Next, the Bragg signal estimates were smoothed across beam. The smoothing was carried out by calculating a running average across beam of Bragg power estimates using a 10 beam window.

The smoothing created a pool of 40 Bragg estimates (10 beam, each with 4 pulses) on which each data point is based. The distribution of these estimates were used to construct the 25th and 75th percentile curves around the data. These curves are roughly +1 and −2 dB relative to the mean data values.

Figure 6 is a geo-plot of the observed field at three fre-

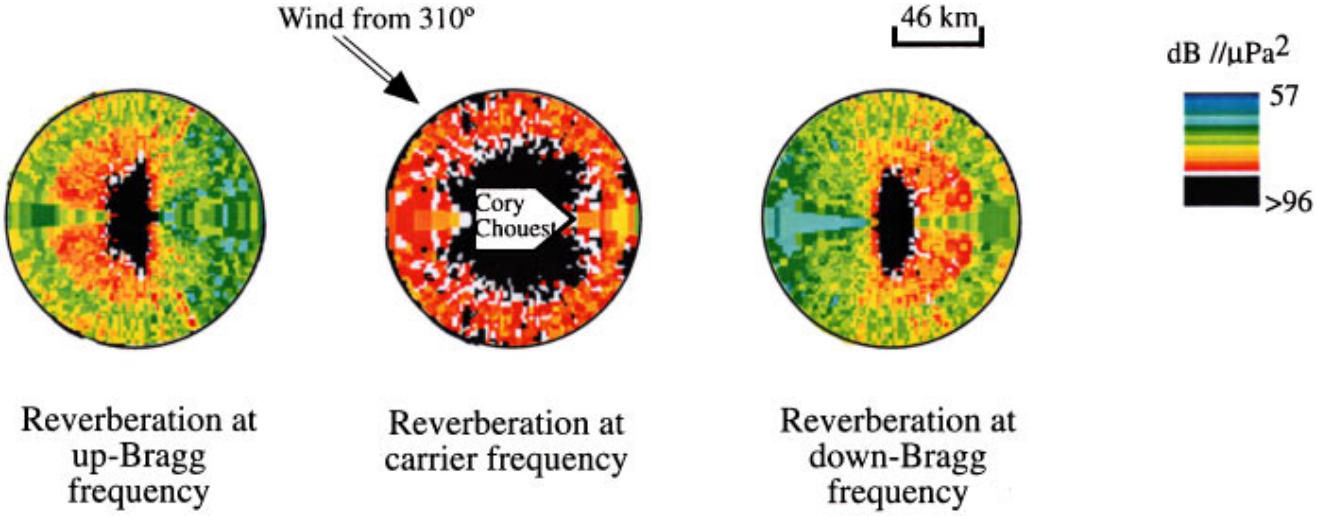


FIG. 6. Spatial directionality of the spectral components of the reverberation field is shown here. The up Bragg lobe is emphasized looking toward the wind, and the down Bragg lobe is emphasized looking away from the wind. The reverberation at the carrier frequency, in contrast, is generally isotropic.

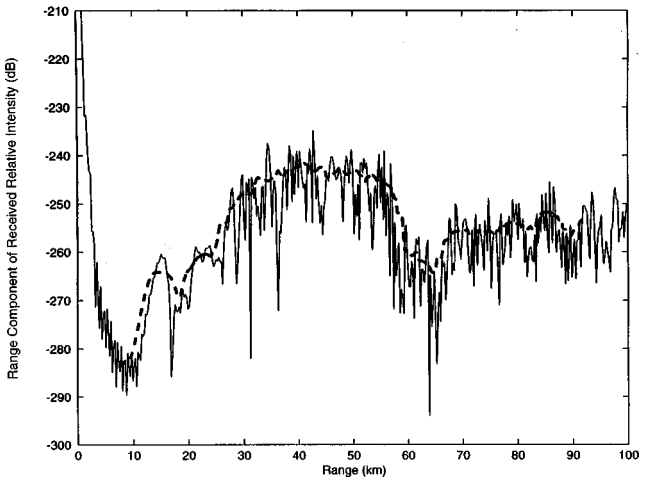


FIG. 7. Modeled received relative intensity versus range for a 10-element source array steered 12 degrees down. The effective source depth is 137 m. The dashed curve represents an average over the range extent of the transmitted pulse.

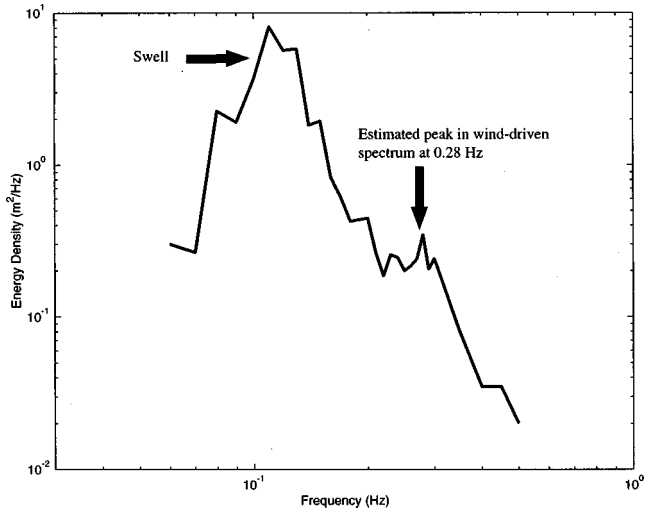


FIG. 8. Gravity wave frequency spectrum measured during Bragg reverberation experiment. Low frequency content is due to residual swell. Seas were developing under a light northwest wind. A peak in the wind-driven part of the spectrum appears at approximately 0.28 Hz.

quencies: up Bragg, down Bragg, and carrier (250 Hz) or ''zero Doppler.'' The viewpoint is that of an observer looking down on the ship. Reverberation level is color coded in dB//μPa^2, according to the scale on the right of the figure. Wind and ship direction are as indicated in the center plot. One can see the directional biases present at the up and down Bragg frequencies, and the azimuthal isotropy present at zero Doppler due to its origin in scattering from bubbles. In the next section the up and down Bragg reverberation levels at the ''pseudo-CZ'' range are calculated, as a function of azimuth using the model described in Sec. I, and compared to the measured reverberation.

III. COMPARISON OF MODELS WITH EXPERIMENTAL DATA

Model inputs were constructed using an *in situ* sound speed profile (measured at the time the Bragg data were taken; Fig. 4), and bottom sound speed and density parameters derived from the Bottom Loss Upgrade Database. A total of 590 propagating modes were computed for the 250 Hz source frequency. Modes 3–590 have equivalent ray grazing angles at the surface between 2° and 24°, corresponding to Bragg frequency shifts in the range $4.38 \leqslant \Omega \leqslant 4.59$ rad/s (0.697 Hz$<$frequency$<$0.731).

The received pressure $P_{\mathrm{rec}}(r)$ was computed using Eq. (21) with the vertical scattering strength $\sigma_v(\theta_n, \theta_m)$. Figure 7 displays the range component of the received relative intensity $\mathrm{RI} = 10 \log_{10} |P_{\mathrm{rec}}(r)|^2$ as a function of range for the modes of interest (that is, ''high angle'' modes have been eliminated, thereby supressing the fathometer returns visible in Fig. 5). Also shown is the intensity averaged in the range window $c\tau/2 \approx 9$ km (the range extent of the transmitted pulse). The relative intensity indicates the round trip transmission loss (TL$=-$RI) between the source, the scattering patch and the receiver. The high-intensity interval between 35 and 50 km is consistent with the time delays observed in the peak Bragg-shifted reverberation data. This range interval corresponds to a region of constructive interference between modes with equivalent ray angles in the range $10° \leqslant \theta_n \leqslant 15°$

A measured gravity wave frequency spectrum[13] from the time period of the Bragg reverberation measurement is shown in Fig. 8. The wind was light (6.8 m/s) during this interval and the seas were developing, with a residual swell component dominating at low frequencies (0.1 Hz). A high frequency peak, due to the newly-shifted northwest wind, appears in the wind-driven part of the spectrum at approximately 0.28 Hz (1.76 rad/s). The value chosen for the peak frequency in the Donelan model for the wind-driven gravity wave component was $\Omega_p = 1.76$ rad/s. The implied phase speed at this peak frequency is $c_p = 5.6$ m/s. The 250 Hz acoustic frequency of the cw pulse yielded a Bragg frequency shift of approximately $\Omega_0 = 4.4$ rad/s (0.7 Hz).

Donelan *et al.* parameterized the spreading parameter β in terms of the dimensionless frequency Ω/Ω_p for frequencies in the range $1 \leqslant \Omega/\Omega_p \leqslant 1.6$. For larger values of β, Donelan's data exhibited an additional dependence on the wave age c_p/U_c. For Critical Sea Test 4 values of these ratios $\Omega/\Omega_p \simeq 2.5$ and $U_c/c_p = 1.2$, the data curves in Fig. 32 of Ref. 6 indicate $\beta \simeq 1.24$.

The azimuth term $H(\phi)$ in Eq. (B9) for the beam intensity was computed using the normalized beam pattern of the horizontal receive array with the spreading function $h(\phi_D, \Omega_0)$ of Eq. (8). An approximate 44-deg difference between the array heading and the wind direction was accounted for in this computation. The range component of the received intensity $|P_{\mathrm{rec}}(r)|^2$ was combined with H according to Eq. (B9) to give the range dependent beam intensity $I(r, \phi)$. Using $I(r, \phi)$ in Eq. (19) with the measured source level and the 3 dB beamwidth function $w(\phi)$ produced final model predictions of the up- and down-Bragg reverberation.

Figure 9(a) and (b) compares the model predictions with the 250 Hz Bragg reverberation data at the up- and down-shifted Bragg frequencies, respectively. The model curves represent an azimuthal slice through $\mathrm{RL}(r, \phi)$ at its peak

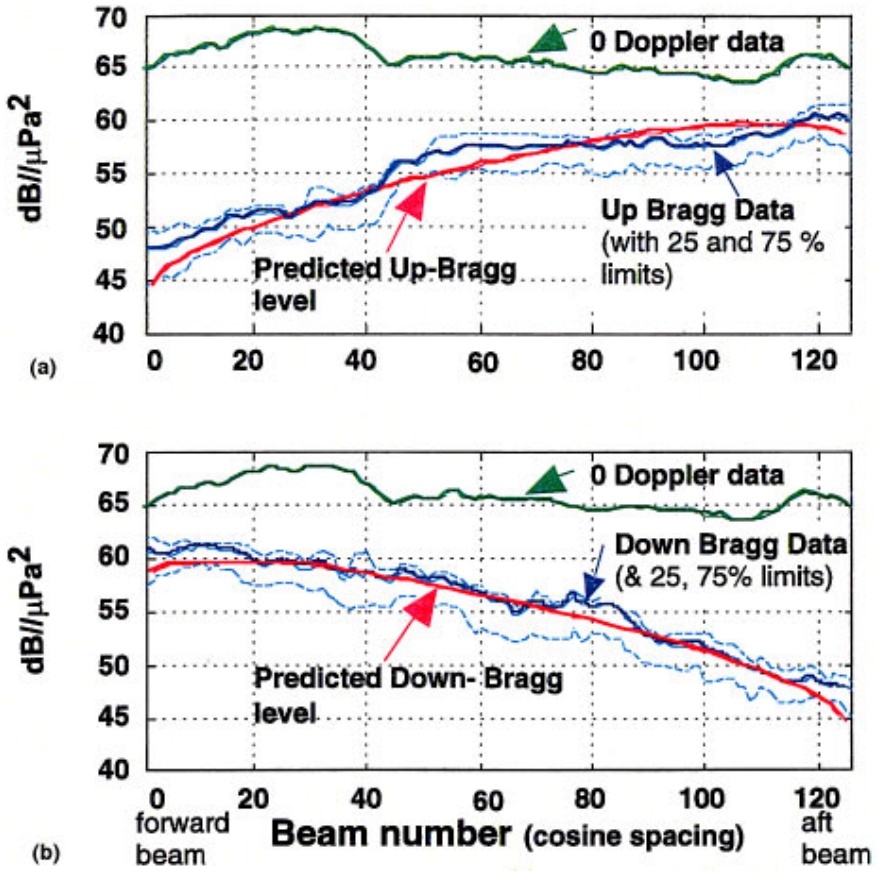


FIG. 9. Predictions of the ''pseudo-CZ'' Bragg scattering levels across azimuth as calculated using the Donelan wave spectrum, first-order perturbation theory, and a normal mode propagation model agree well with the observed Bragg levels extracted from each beam of the horizontal line array. In contrast the zero Doppler shifted data is nearly isotropic indicating that it is not rough-surface related. (The water wave spectrum parameter $\beta = 1.24$.) (a) Up Bragg data; (b) Down Bragg data.

range, consistent with the measured values. The underlying range interval for these predictions is approximately 37–46 km. The green data curve in each figure represents the measured zero-Doppler reverberation. The solid black curves represent the mean level of the Bragg-shifted data, and the dashed blue curves indicate the 25th and 75th percentiles of the data on which the averages were based. The model predictions (magenta curves) are seen to agree very well with the data in both absolute level and azimuthal variation. The predicted levels are generally confined between the 25th and 75th percentile curves.

The model predictions of Fig. 9(a) and (b) are based upon estimates of the underlying physical parameters. While uncertainties exist in these parameter values, a cursory examination of the effects of small variations in wind speed (± 2 m/s), wind direction ($\pm 5°$), and peak water wave frequency (± 0.02 Hz) revealed only minor differences ($\simeq \pm 1$ dB) in the predicted levels. It is concluded that the heuristic normal mode reverberation model combined with the 2-D Donelan water wave spectrum model and first-order perturbation scattering theory accounts reasonably well for the levels and azimuthal variations observed in the Critical Sea Test 4 reverberation measurements.

IV. SUMMARY

We have extended the theory of acoustic surface scattering to include scattering from a moving sea surface described by Donelan's asymmetric water wave spectrum. Using environmental data from Critical Sea Test 4 (CST-4), we computed a matrix of acoustic scattering coefficients for all significant combinations of up and down going incident and surface scattered equivalent ray angles. This scattering matrix, in conjunction with a normal mode scattering model, a beam pattern model, and the sound speed profile for the CST-4 experiment yielded a prediction for the up-Bragg and down-Bragg reverberation as a function of azimuth around the receiver. This prediction agreed well with the actual Bragg-shifted reverberation measured during the CST-4 experiment.

From these results we conclude:

(1) Surface scattered acoustic waves show a distinct azimuthal dependence governed by the water wave spectrum and the spatial filtering characteristics of the receiver (in our case, a linear array);
(2) Acoustic backscatter amplitudes are well predicted by first-order perturbation theory applied via the equivalent angles of incoming and outgoing normal modes;
(3) The water wave spectrum proposed by Donelan *et al.*[6] leads to predictions for acoustic backscatter which agree with the measurements both in azimuthal dependence and acoustic backscatter intensity.

ACKNOWLEDGMENTS

The authors thank Dr. E. I. Thorsos (APL/UW) for his helpful suggestions and gratefully acknowledge the support of the U. S. Space and Naval Warfare Systems Command (PMW-182).

APPENDIX A

1. Fourier integral solution of the wave equation

The wave equation for a homogeneous medium is

$$\nabla^2\psi(\mathbf{r},t)-\frac{1}{c^2}\frac{\partial^2\psi(\mathbf{r},t)}{\partial t^2}=0, \tag{A1}$$

where $\psi(\mathbf{r},t)$ is the acoustic pressure field and c is the speed of sound. We can resolve $\psi(\mathbf{r},t)$ in terms of a Fourier integral

$$\psi(\mathbf{r},t)=\int_{-\infty}^{\infty}\Psi(\mathbf{r},\omega)e^{-i\omega t}\,d\omega, \tag{A2}$$

where $\Psi(\mathbf{r},\omega)$ is the Fourier amplitude which satisfies the Helmholtz equation, and $\omega/2\pi$ is the frequency of the acoustic wave.

Now let

$$\mathbf{K}=K_x\hat{e}_x+K_y\hat{e}_y \tag{A3}$$

and

$$\boldsymbol{\kappa}=\kappa_x\hat{e}_x+\kappa_y\hat{e}_y+\kappa_z\hat{e}_z=\mathbf{K}+\kappa_z\hat{e}_z. \tag{A4}$$

Here $\hat{e}_x$, $\hat{e}_y$ and $\hat{e}_z$ are the unit base vectors along the x, y, and z axes, respectively. Further let

$$\mathbf{R}=x\hat{e}_x+y\hat{e}_y \tag{A5}$$

and

$$\mathbf{r}=x\hat{e}_x+y\hat{e}_y+z\hat{e}_z=\mathbf{R}+z\hat{e}_z. \tag{A6}$$

Now a well known representation for $\Psi(\mathbf{r},\omega)$ as a two-dimensional Fourier integral is

$$\Psi(\mathbf{r},\omega)=\int_{-\infty}^{\infty}d^2K\{\phi(\mathbf{K},\omega)e^{\pm i\beta(K,\omega)z}\}e^{i\mathbf{K}\cdot\mathbf{R}}, \tag{A7}$$

where $d^2K=dK_x\,dK_y$, and

$$\beta(K,\omega)=+\sqrt{\frac{\omega^2}{c^2}-K^2}. \tag{A8}$$

In Eq. (A7) the ''+'' sign is used for the scattered wave and the ''−'' sign for the incident wave. We have taken the z-axis positive downward.

Combining these results we get

$$\psi(\mathbf{r},t)=\int_{-\infty}^{\infty}d^2K\,d\omega\,\phi(\mathbf{K},\omega)e^{i[\mathbf{K}\cdot\mathbf{R}\pm\beta(K,\omega)z-\omega t]}. \tag{A9}$$

2. T-matrix

We need a scattering operator which transforms the incident field on an object into the scattered field. For the rough surface case it is convenient to work with the Fourier amplitudes. We denote the scattering operator in $(\mathbf{K},\omega)$ space by T and define it as

$$\phi^s(\mathbf{K}_s,\omega_s)=\int d^2K\,d\omega\,T(\mathbf{K}_s,\omega_s,\mathbf{K}_i,\omega_i)\phi^i(\mathbf{K}_i,\omega_i), \tag{A10}$$

where ϕ^s = amplitude of the scattered wave, ϕ^i = amplitude of the incident wave, $\mathbf{K}_s$ = horizontal component of the scattered wave number vector, $\mathbf{K}_i$ horizontal component of the incident wave number vector, ω_s = frequency of scattered wave, and ω_i frequency of incident wave. Thus, the operator T connects the incident field to the scattered field in $(\mathbf{K},\omega)$ space.

We will use the Rayleigh hypothesis and perturbation theory. Rayleigh considered the problem of a sinusoidal surface bounding a homogeneous half-space and being irradiated by an incident plane wave. Rayleigh assumed that the scattered wave could be represented everywhere in the homogeneous half-space by a linear combination of plane waves each of which propagates or is evanescent away from the surface.

The Rayleigh hypothesis goes one step further and assumes that the representation given by Eq. (A9) for the scattered wave is valid for every point above the surface, even for those points in the troughs. This assumption has been questioned because in this region one might expect to find waves traveling toward the surface and Eq. (A9) describes only the wave moving away from the surface or decaying exponentially in the direction away from the surface. In Ref. 8 the range of validity of the Rayleigh hypothesis is discussed. For our case it is a valid hypothesis.

We now proceed with the perturbation solution. Let the surface be given in the form

$$z=\bar{h}h(\mathbf{R},t), \tag{A11}$$

where $\bar{h}$ is the rms surface height.

We expand the T operator in powers of $\bar{h}$ as follows:

$$T(\mathbf{K}_s,\omega_s,\mathbf{K}_i,\omega_i)=\sum_{n=0}^{\infty}\frac{(\bar{h})^n}{n!}T^{(n)}(\mathbf{K}_s,\omega_s,\mathbf{K}_i,\omega_i). \tag{A12}$$

The boundary condition requires that the total field vanish on the surface:

$$\int_{-\infty}^{\infty}d^2K_i\,d\omega_i\phi^i(\mathbf{K}_i,\omega_i)e^{i[\mathbf{K}_i\cdot\mathbf{R}-\beta(K_i,\omega_i)\bar{h}h-\omega_i t]}+\int_{-\infty}^{\infty}d^2K_s'\,d\omega_s'\phi^s(\mathbf{K}_s',\omega_s')e^{i[\mathbf{K}_s'\cdot\mathbf{R}+\beta(K_s',\omega_s')\bar{h}h-\omega_s' t]}=0. \tag{A13}$$

First we multiply Eq. (A13) by $\exp\{-i[\mathbf{K}_s\cdot\mathbf{R}-\omega_s t]\}$ and integrate with respect to $(\mathbf{R},t)$. Then using Eq. (A10), as well as the perturbation expansion for T given by Eq. (A12), we expand the exponential $\exp\{-i\beta(K_i,\omega_i)\bar{h}h\}$ on the left side of the equation and the exponential $\exp\{i\beta(K_s',\omega_s')\bar{h}h\}$ on the right side of the equation. After a lengthy calculation we can equate corresponding powers of $\bar{h}$. Details of this calculation can be found in Refs. 8 and 10. This gives for the zeroth order

$$T^{(0)}(\mathbf{K}_s,\omega_s;\mathbf{K}_i,\omega_i)=-\delta(\mathbf{K}_s-\mathbf{K}_i)\,\delta(\omega_s-\omega_i). \tag{A14}$$

This is the specularly reflected wave. Note that since this equation implies $\omega_s=\omega_i$, the specularly reflected wave ex-

periences no Doppler shift due to the moving sea surface.

The first-order term in the expansion for the T-matrix is

$$T^{(1)}(\mathbf{K}_s,\omega_s;\mathbf{K}_i,\omega_i)=2i\beta(\mathbf{K}_i,\omega_i)\tilde{h}(\mathbf{K}_s-\mathbf{K}_i,\omega_s-\omega_i), \tag{A15}$$

where we have defined

$$\tilde{h}(\mathbf{K},\omega)=\frac{1}{(2\pi)^3}\int_{-\infty}^{\infty} d^2R\,dt\,h(\mathbf{R},t)e^{-i(\mathbf{K}\cdot\mathbf{R}-\omega t)}. \tag{A16}$$

3. Incident wave field

In Sec. II we have derived the T-matrix for an infinite surface. The scattering strength is defined in terms of a far-field scattered intensity from a finite surface area. If we used an incident plane wave with an infinite rough surface, then we could not talk about a far-field intensity let alone scattering from a finite surface area. Conceptually, we could taper surface roughness (and throw away all specular scattering including that from the flat edges), and keep the incoming plane wave; then keeping the tapering way out on the edges making edge effects negligible, and calculate cross section per unit area. However, it is more practical to use an incident plane wave that is spatially tapered. Thorsos[14] derived such a solution for a one-dimensional surface. We will extend his procedure to a two-dimensional surface.

We start with a solution of the wave equation that is constructed from an angular spectrum of plane waves at a single frequency. Using the coordinate system shown in Fig. 1, we can write

$$\begin{aligned}\psi^i(\mathbf{r},t)=&\frac{e^{-i\omega_i t}}{2\pi\sigma_\theta\sigma_\phi}\int_{-\pi}^{\pi}d\phi'\int_{-\pi/2}^{\pi/2}d\theta'\\&\times e^{-(\theta'-\theta_i)^2/2\sigma_\theta^2-(\phi'-\phi_i)^2/2\sigma_\phi^2}\\&\times e^{i(\kappa_i x\cos\theta'\cos\phi'+\kappa_i y\cos\theta'\sin\phi'-\kappa_i z\sin\theta')}.\end{aligned} \tag{A17}$$

For the incident wave

$$\kappa_i\equiv\frac{\omega_i}{c}. \tag{A18}$$

The parameters σ_θ and σ_ϕ control the tapering of the incident wave field. The maximum pressure is in the direction (θ_i,ϕ_i). The incident wave representation given by Eq. (A17) is too complicated to use. We need to use an approximate expression. To this end we let $\delta_\theta=\theta'-\theta_i$ and $\delta_\phi=\phi'-\phi_i$. We expand

$$\begin{aligned}&\kappa_i x\cos(\delta_\theta+\theta_i)\cos(\delta_\phi+\phi_i)\\&\quad+\kappa_i y\cos(\delta_\theta+\theta_i)\sin(\delta_\phi+\phi_i)-\kappa_i z\sin(\delta_\theta+\theta_i)\end{aligned}$$

in δ_θ and δ_ϕ, keeping terms to second order. Then we transform the integration variables to δ_θ and δ_ϕ, and extend the limits of integration to $\pm\infty$. Upon carrying out the integrations and keeping only second order terms, we get after an extremely lengthy calculation

$$\begin{aligned}\psi^i(\mathbf{r},t)=&e^{-i\omega_i t}e^{i(\kappa_{ix}x+\kappa_{iy}y-\kappa_{iz}z)}e^{-(\sigma_\phi^2/2)(\alpha_x x+\alpha_y y)^2}\\&\times e^{-(\sigma_\theta^2/2)(\beta_x x+\beta_y y+\beta_z z)^2},\end{aligned} \tag{A19}$$

where we have put

$$\begin{aligned}&\kappa_{ix}=\kappa_i\cos\theta_i\cos\phi_i,\\&\kappa_{iy}=\kappa_i\cos\theta_i\sin\phi_i,\\&\kappa_{iz}=\kappa_i\sin\theta_i,\\&\alpha_x=\kappa_{iy},\quad \alpha_y=-\kappa_{ix},\\&\beta_x=\kappa_{iz}\cos\phi_i,\quad \beta_y=\kappa_{iz}\sin\phi_i,\quad \beta_z=\kappa_i\cos\theta_i,\end{aligned} \tag{A20}$$

We now need to determine $\phi^i(\mathbf{K},\omega)$. Taking the inverse transform of Eq. (A9) we get

$$\phi^i(\mathbf{K},\omega)e^{-i\beta(K,\omega)z}=\frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}d^2R\,dt\,\psi^i(\mathbf{r},t)e^{-i(\mathbf{K}\cdot\mathbf{R}-\omega t)}. \tag{A21}$$

Using Eq. (A19) and keeping only second-order terms, we get upon carrying out the integration

$$\phi^i(\mathbf{K},\omega)=\frac{\delta(\omega_i-\omega)}{4\pi p_x p_y}e^{-(1/4p_x^2)(\kappa_{ix}-K_x)^2-(1/4p_y^2)(\kappa_{iy}-K_y)^2}, \tag{A22}$$

where we have put

$$p_x^2\equiv\frac{\sigma_\phi^2}{2}\alpha_x^2+\frac{\sigma_\theta^2}{2}\beta_x^2,\quad p_y^2\equiv\frac{\sigma_\phi^2}{2}\alpha_y^2+\frac{\sigma_\theta^2}{2}\beta_y^2. \tag{A23}$$

4. The scattered wave

In order to calculate the surface scattering strength, we need an expression for the scattered pressure in the farfield. We start with the representation given by Eq. (A7),

$$\Psi^s(\mathbf{r},\omega_s)=\int_{-\infty}^{\infty}dK_{sx}\,dK_{sy}\phi^s(\mathbf{K}_s,\omega_s)e^{i\kappa_s r\cos\gamma}. \tag{A24}$$

Here

$$\cos\gamma=\frac{\boldsymbol{\kappa}_s\cdot\mathbf{r}}{\kappa_s r}=\frac{K_{sx}x+K_{sy}y+\kappa_{sz}z}{\kappa_s r}. \tag{A25}$$

Also note that κ_{sz} is a function of K_{sx} and K_{sy} through the relation

$$\kappa_{sz}^2=\kappa_s^2-K_{sx}^2-K_{sy}^2. \tag{A26}$$

We will use the method of stationary phase to evaluate Eq. (A24) in the farfield. First, let us list the result of applying the method of stationary phase to the general integral

$$I(\epsilon)=\int_{-\infty}^{\infty}du_1\,du_2\,g(\mathbf{u})e^{i\epsilon f(\mathbf{u})},\quad \mathbf{u}=(u_1,u_2). \tag{A27}$$

The first-order term in the stationary phase expansion of Eq. (A27) is[15]

$$I_0(\epsilon)\approx\frac{2\pi}{\epsilon}\frac{g(\mathbf{u}_0)\exp\left\{i\epsilon f(\mathbf{u}_0)+\frac{\pi i}{4}\,\mathrm{sig}(F(\mathbf{u}_0))\right\}}{\sqrt{\det F(\mathbf{u}_0)}}. \tag{A28}$$

Here $\mathbf{u}_0$ is the stationary point, and F is the matrix

$$F=\begin{bmatrix} \dfrac{\partial^2 f}{\partial u_1^2} & \dfrac{\partial^2 f}{\partial u_1\,\partial u_2} \\ \dfrac{\partial^2 f}{\partial u_2\,\partial u_1} & \dfrac{\partial^2 f}{\partial u_2^2} \end{bmatrix}. \tag{A29}$$

$\mathrm{Det}\,F$ is the determinant of the matrix F, and $\mathrm{sig}\,F$ is the signature of the matrix F. The signature of F is the number of positive eigenvalues minus the number of negative eigenvalues. In this case $\mathrm{sig}\,F=-2$.

Comparing Eqs. (A24) and (A27), we identify $(u_1,u_2)=(K_{sx},K_{sy})$, $g=\phi^s$, $\epsilon=r$, and $f=\kappa_s\cos\gamma$. The stationary point is the solution of the equations

$$\frac{\partial f}{\partial K_{sx}}=\frac{\partial f}{\partial K_{sy}}=0. \tag{A30}$$

Using Eq. (A26), we note

$$\frac{\partial \kappa_{sz}}{\partial K_{sx}}=-\frac{K_{sx}}{\kappa_{sz}},\quad \frac{\partial \kappa_{sz}}{\partial K_{sy}}=-\frac{K_{sy}}{\kappa_{sz}}. \tag{A31}$$

We also note

$$\kappa_{sz}=\kappa_s\sin\theta=\kappa_s\frac{z}{r}. \tag{A32}$$

Thus, Eqs. (A30) becomes

$$\begin{aligned} \frac{\partial f}{\partial K_{sx}}&=\frac{x}{r}+\frac{z}{r}\frac{\partial \kappa_{sz}}{\partial K_{sx}}=0,\\ \frac{\partial f}{\partial K_{sy}}&=\frac{y}{r}+\frac{z}{r}\frac{\partial \kappa_{sz}}{\partial K_{sy}}=0. \end{aligned} \tag{A33}$$

Using Eq. (A31), we get for the stationary points

$$K_{sx}=\kappa_s\frac{x}{r},\quad K_{sy}=\kappa_s\frac{y}{r}. \tag{A34}$$

It is easy to show that

$$\det F(\mathbf{u}_0)=\frac{1}{\kappa_{sz}^2}. \tag{A35}$$

Putting all of these results together, we get from Eq. (A29) for our case

$$\Psi^s(\mathbf{r},\omega_s)\approx 2\pi\kappa_{sz}\phi^s(\mathbf{K}_s,\omega_s)e^{i(\pi/4)\ \mathrm{sig}\ F}\frac{e^{i\kappa_s r}}{r}. \tag{A36}$$

5. Asymmetric sea surface wave spectrum

The equation describing a time varying sea surface is given by Eq. (A11). If we assume that the wave process is homogeneous and stationary, the autocorrelation function $C(\mathbf{R},t)$ is

$$C(\boldsymbol{\xi},\tau)=\overline{h^2}\langle h(\mathbf{R},t)h(\mathbf{R}+\boldsymbol{\xi},t+\tau)\rangle, \tag{A37}$$

where $\langle\ \rangle$ denotes an ensemble average.

The wave spectrum is defined to be the Fourier transform of the autocorrelation function, so it is given by

$$W(\boldsymbol{\Gamma},\Omega)=\frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}d^2\xi\,d\tau C(\boldsymbol{\xi},\tau)e^{-i(\boldsymbol{\Gamma}\cdot\boldsymbol{\xi}-\Omega\tau)}, \tag{A38}$$

where $\boldsymbol{\Gamma}$ is the wave number vector and $\Omega/2\pi$ is the frequency of the water waves. Consequently, $C(\boldsymbol{\xi},t)$ is given by the inverse transform

$$C(\boldsymbol{\xi},\tau)=\int_{-\infty}^{\infty}d^2\Gamma\,d\Omega\,W(\boldsymbol{\Gamma},\Omega)e^{i(\boldsymbol{\Gamma}\cdot\boldsymbol{\xi}-\Omega\tau)}. \tag{A39}$$

Because $h(\mathbf{R},t)$ is a real function, and because the autocorrelation function has the property that

$$C(\boldsymbol{\xi},\tau)=C(-\boldsymbol{\xi},-\tau), \tag{A40}$$

the wave spectrum is a real function, and has the symmetry property

$$W(\boldsymbol{\Gamma},\Omega)=W(-\boldsymbol{\Gamma},-\Omega). \tag{A41}$$

Using Eq. (A41), we can rewrite Eq. (A39) in the following manner

$$\begin{aligned} C(\boldsymbol{\xi},\tau)=\frac{1}{2}\int_{-\infty}^{\infty}d^2\Gamma\,d\Omega\{&W(\boldsymbol{\Gamma},\Omega)\\ &+W(-\boldsymbol{\Gamma},-\Omega)\}\,e^{i(\boldsymbol{\Gamma}\cdot\boldsymbol{\xi}-\omega\tau)}. \end{aligned} \tag{A42}$$

We will assume that the following dispersion relation holds for the water waves

$$\Omega=\Omega_\Gamma=\sqrt{g|\boldsymbol{\Gamma}|}, \tag{A43}$$

where g is the acceleration of gravity. Then we can write Eq. (A42) as

$$\begin{aligned} C(\boldsymbol{\xi},\tau)=\int_{-\infty}^{\infty}d^2\Gamma\,d\Omega e^{i(\boldsymbol{\Gamma}\cdot\boldsymbol{\xi}-\Omega\tau)}\\ \times\{\tfrac{1}{2}[W(\boldsymbol{\Gamma})\,\delta(\Omega-\Omega_\Gamma)+W(-\boldsymbol{\Gamma})\,\delta(\Omega+\Omega_\Gamma)]\}. \end{aligned} \tag{A44}$$

6. Scattering strength

Following Urick,[16] we define the scattering strength S as

$$S=10\,\mathrm{Log}\,\sigma, \tag{A45}$$

where the scattering coefficient σ is

$$\sigma=\frac{\langle I_s\rangle r^2}{I_i A}. \tag{A46}$$

In Eq. (A46) I_i is the intensity of the incident wave, A is the ensonified area of the surface, $\langle I_s\rangle$ is the average scattered intensity at the receiver, and r is the distance from the surface to the receiver.

The power E incident on the surface is

$$E=I_i A\sin\theta_i, \tag{A47}$$

so we can write Eq. (A46) as

$$\sigma=\frac{\langle I_s\rangle r^2\sin\theta_i}{E}. \tag{A48}$$

Because the incident wave is tapered, the incident power is not uniform over the surface. Let dE be the power on a small area $dxdy$. Then from Eq. (A47)

$$dE=\frac{\sin\theta_i}{2\rho c}|\psi^i(x,y,z=0,t)|^2dx\,dy, \tag{A49}$$

where ρ is the density of water, and we have used Eq. (A19) for the pressure in the incident wave.

The total power incident on the surface is then

$$E=\frac{\sin\theta_i}{2\rho c}\int_{-\infty}^{\infty}dx\,dy e^{-\sigma_\phi^2(\alpha_x x+\alpha_y y)^2}e^{-\sigma_\theta^2(\beta_x x+\beta_y y)^2}$$

$$\approx\frac{\pi\sin\theta_i}{4\rho c p_x p_y}, \tag{A50}$$

where p_x and p_y are given by Eq. (A23). In the final result in Eq. (A50), we neglected the cross terms xy in the exponents. After integration these cross terms give rise to terms which contain powers of σ_θ and σ_ϕ higher than two. Now

$$\langle I_s\rangle=\frac{\langle|\Psi^s(\mathbf{r},\omega_s)|^2\rangle}{2\rho c}, \tag{A51}$$

where Ψ^s is given by Eq. (A36). Thus, we can write σ in the form

$$\sigma=8\pi p_x p_y\kappa_{sz}^2\langle\phi^s(\mathbf{K}_s,\omega_s)\phi^s{*}(\mathbf{K}_s,\omega_s)\rangle, \tag{A52}$$

where the $*$ denotes the complex conjugate.

Now from Eqs. (A10) and (A22), we get

$$\phi^s(\mathbf{K}_s,\omega_s)=\frac{1}{4\pi p_x p_y}\int_{-\infty}^{\infty}dK_x\,dK_y\,d\omega$$
$$\times T(\mathbf{K}_s,\omega_s;\mathbf{K},\omega)\,\delta(\omega_i-\omega)$$
$$\times e^{-(1/4p_x^2)(\kappa_{ix}-K_x)^2-(1/4p_y^2)(\kappa_{iy}-K_y)^2}. \tag{A53}$$

After carrying out the integration over ω, we get

$$\langle\phi^s(\mathbf{K}_s,\omega_s)\phi^s{*}(\mathbf{K}_s,\omega_s)\rangle$$
$$=\frac{1}{(4\pi p_x p_y)^2}\int_{-\infty}^{\infty}dK_{1x}\,dK_{1y}$$
$$\times e^{-(1/4p_x^2)(\kappa_{ix}-K_{1x})^2-(1/4p_y^2)(\kappa_{iy}-K_{1y})^2}\int_{-\infty}^{\infty}dK_{2x}\,dK_{2y}$$
$$\times e^{-(1/4p_x^2)(\kappa_{ix}-K_{2x})^2-(1/4p_y^2)(\kappa_{iy}-K_{2y})^2}$$
$$\times\langle T(\mathbf{K}_s,\omega_s;\mathbf{K}_1\omega_i)T^*(\mathbf{K}_s,\omega_s;\mathbf{K}_2,\omega_i)\rangle. \tag{A54}$$

To calculate $\langle TT*\rangle$, we will use only the first-order contributon to the T-matrix given by Eq. (A15). Thus,

$$\langle T(\mathbf{K}_s,\omega_s;\mathbf{K}_1,\omega_1)T*(\mathbf{K}_s,\omega_s;\mathbf{K}_2,\omega_2)\rangle$$
$$=\overline{h^2}\langle T^{(1)}(\mathbf{K}_s,\omega_s;\mathbf{K}_1,\omega_1)T^{(1)}*(\mathbf{K}_s,\omega_s;\mathbf{K}_2,\omega_2)\rangle$$
$$=4\overline{h^2}\beta(K_1,\omega_1)\beta(K_2,\omega_2)\langle\tilde{h}(\mathbf{K}_s-\mathbf{K}_1,\omega_s-\omega_1)$$
$$\times\tilde{h}*(\mathbf{K}_s-\mathbf{K}_2,\omega_s-\omega_2)\rangle. \tag{A55}$$

Now using Eq. (A16)

$$\overline{h^2}\langle\tilde{h}(\mathbf{K}_s-\mathbf{K}_1,\omega_s-\omega_1)\tilde{h}*(\mathbf{K}_s-\mathbf{K}_2,\omega_s-\omega_2)\rangle$$
$$=\frac{1}{(2\pi)^6}\int_{-\infty}^{\infty}d^2R_1\,dt_1\,d^2R_2\,dt_2\,C(\mathbf{R}_1,t_1;\mathbf{R}_2,t_2)$$
$$\times e^{-i[(\mathbf{K}_s-\mathbf{K}_1)\cdot\mathbf{R}_1-(\omega_s-\omega_1)t_1]}e^{i[(\mathbf{K}_s-\mathbf{K}_2)\cdot\mathbf{R}_2-(\omega_s-\omega_2)t_2]}, \tag{A56}$$

where $C(\mathbf{R}_1,t_1;\mathbf{R}_2,t_2)$ is the autocorrelation function defined by Eq. (A37). Let us make the following change of variables in order to integrate Eq. (A56):

$$\boldsymbol{\xi}=\mathbf{R}_1-\mathbf{R}_2,\quad \boldsymbol{\eta}=\tfrac{1}{2}(\mathbf{R}_1+\mathbf{R}_2),$$
$$\tau=t_1-t_2,\quad \alpha=\tfrac{1}{2}(t_2+t_1). \tag{A57}$$

Then Eq. (A56) becomes

$$\overline{h^2}\langle\tilde{h}(\mathbf{K}_s-\mathbf{K}_1,\omega_s-\omega_1)\tilde{h}*(\mathbf{K}_s-\mathbf{K}_2,\omega_s-\omega_2)\rangle$$
$$=\frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}d^2\eta\,d\alpha e^{-i(\mathbf{K}_2-\mathbf{K}_1)\cdot\boldsymbol{\eta}}e^{i(\omega_2-\omega_1)\alpha}$$
$$\times\frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}d^2\xi\,d\tau C(\boldsymbol{\xi},\tau)$$
$$\times e^{-i[\mathbf{K}_s-(\mathbf{K}_1/2)-(\mathbf{K}_2/2)]\cdot\boldsymbol{\xi}}e^{i[\omega_s-(\omega_2/2)-(\omega_1/2)]\tau}. \tag{A58}$$

Upon carrying out the integrations, we get

$$\overline{h^2}\langle\tilde{h}(\mathbf{K}_s-\mathbf{K}_1,\omega_s-\omega_1)\tilde{h}*(\mathbf{K}_s-\mathbf{K}_2,\omega_s-\omega_2)\rangle$$
$$=\delta(\mathbf{K}_2-\mathbf{K}_1)\,\delta(\omega_2-\omega_1)$$
$$\times\overline{W}\left(\mathbf{K}_s-\frac{\mathbf{K}_1}{2}-\frac{\mathbf{K}_2}{2},\omega_s-\frac{\omega_2}{2}-\frac{\omega_1}{2}\right), \tag{A59}$$

where we have put

$$\overline{W}(\boldsymbol{\Gamma},\Omega)=\tfrac{1}{2}[W(\boldsymbol{\Gamma})\delta(\Omega-\Omega_\Gamma)+W(-\boldsymbol{\Gamma})\delta(\Omega+\Omega_\Gamma)] \tag{A60}$$

from Eq. (A44). Combining Eqs. (A59) and (A55), we get

$$\langle T(\mathbf{K}_s,\omega_s;\mathbf{K}_1,\omega_1)T*(\mathbf{K}_s,\omega_s;\mathbf{K}_2,\omega_2)\rangle$$
$$=4\beta(K_1,\omega_1)\beta(K_2,\omega_2)\,\delta(\mathbf{K}_2-\mathbf{K}_1)\,\delta(\omega_2-\omega_1)$$
$$\times\overline{W}\left(\mathbf{K}_s-\frac{\mathbf{K}_1}{2}-\frac{\mathbf{K}_2}{2},\omega_s-\frac{\omega_2}{2}-\frac{\omega_1}{2}\right). \tag{A61}$$

Now note that Eq. (A61) states that $\omega_2=\omega_1$ which we already knew from Eq. (A54). In fact, we see from Eq. (A54) that $\omega_1=\omega_2=\omega_i$. Consequently we can write Eq. (A61) as

$$\langle T(\mathbf{K}_s,\omega_s;\mathbf{K}_1,\omega_i)T*(\mathbf{K}_s,\omega_s;\mathbf{K}_2,\omega_i)\rangle$$
$$=4\beta(K_1,\omega_i)\beta(K_2,\omega_i)\,\delta(\mathbf{K}_2-\mathbf{K}_1)$$
$$\times\overline{W}\left(\mathbf{K}_s-\frac{\mathbf{K}_1}{2}-\frac{\mathbf{K}_2}{2},\omega_s-\omega_i\right). \tag{A62}$$

We now combine Eq. (A62) with Eq. (A54) to get

$$\langle \phi^s(\mathbf{K}_s,\omega_s)\phi^{s*}(\mathbf{K}_s,\omega_s)\rangle = \frac{1}{(4\pi p_x p_y)^2}\int_{-\infty}^{\infty} dK_{1x}\,dK_{1y} e^{-(1/4p_x^2)(\kappa_{ix}-K_{!x})^2-(1/4p_y^2)(\kappa_{iy}-K_{1y})^2}$$
$$\times \int_{-\infty}^{\infty} dK_{2x}\,dK_{2y} e^{-(1/4p_x^2)(\kappa_{ix}-K_{2x})^2-(1/4p_y^2)(\kappa_{iy}-K_{2y})^2} 4\beta(\mathbf{K}_1,\omega_i)\beta(\mathbf{K}_2,\omega_i)\delta(\mathbf{K}_2-\mathbf{K}_1)$$
$$\times \overline{W}\left(\mathbf{K}_s-\frac{\mathbf{K}_1}{2}-\frac{\mathbf{K}_2}{2},\omega_s-\omega_i\right). \tag{A63}$$

Carrying out the integration with respect to $\mathbf{K}_2$ results in

$$\langle \phi^s(\mathbf{K}_s,\omega_s)\phi^{s*}(\mathbf{K}_s,\omega_s)\rangle$$
$$= \frac{1}{(4\pi p_x p_y)^2}\int_{-\infty}^{\infty} dK_{1x}\,dK_{1y}$$
$$\times e^{-(1/2p_x^2)(\kappa_{ix}-K_{1x})^2-(1/2p_y^2)(\kappa_{iy}-K_{1y})^2}$$
$$\times 4\beta^2(\mathbf{K}_1,\omega_i)\overline{W}(\mathbf{K}_s-\mathbf{K}_1,\omega_s-\omega_i). \tag{A64}$$

As σ_θ and σ_ϕ become very small, we may make the approximation

$$\beta^2(\mathbf{K}_1,\omega_i)\overline{W}(\mathbf{K}_s-\mathbf{K}_1,\omega_s-\omega_i)\approx K_{iz}^2\overline{W}(\mathbf{K}_s-\mathbf{K}_i,\omega_s-\omega_i). \tag{A65}$$

Inserting this expression into Eq. (A64) and carrying out the integration with respect to $\mathbf{K}_1$ yields

$$\langle \phi^s(\mathbf{K}_s,\omega_s)\phi^{s*}(\mathbf{K}_s,\omega_s)\rangle$$
$$= \frac{1}{2\pi p_x p_y}\kappa_{iz}^2\,\overline{W}(\mathbf{K}_s-\mathbf{K}_i,\omega_s-\omega_i). \tag{A66}$$

Finally inserting this last expression into Eq. (A52), gives the desired scattering strength

$$\sigma = 2\kappa_{iz}^2\kappa_{sz}^2\{W(\mathbf{K}_s-\mathbf{K}_i)\,\delta(\omega_s-\omega_i-\Omega_\Gamma)$$
$$+ W(-\mathbf{K}_s+\mathbf{K}_i)\,\delta(\omega_s-\omega_i+\Omega_\Gamma)\}. \tag{A67}$$

APPENDIX B: NORMAL MODE REVERBERATION MODEL

The far-field normal mode expression for the complex pressure due to a point source of unit strength is

$$P(r,z) = (2i\pi)^{1/2}\rho^{-1}\sum_{n=1}^{\infty}\frac{1}{\sqrt{k_n r}}\Phi_n(z_s)\Phi_n(z)e^{(ik_n-\delta_n)r}, \tag{B1}$$

where ρ is the density, $\Phi_n(z)$ are the normalized eigenfunctions, k_n are the eigenvalues, δ_n are the modal attenuations, and z_s is the source depth. A harmonic time dependence $e^{-i\omega_s t}$ is assumed. In K. J. McCann's model, the incident field consists of contributions from the up-going parts of the modes evaluated at the surface. The scattered field is comprised of the down-going mode components. The up- and down-going waves $\Phi_n^+(z)$ and $\Phi_n^-(z)$ are found by approximating each mode by its WKB asymptotic expansion near the surface:

$$\Phi_n(z) = \Phi_n^+(z)+\Phi_n^-(z) = \frac{A}{\sqrt{\gamma_n(z)}}\left[\exp\left(i\int^z \gamma_n(\zeta)d\zeta\right) - \exp\left(-i\int^z \gamma_n(\zeta)d\zeta\right)\right], \tag{B2}$$

where $\gamma_n(z)=\sqrt{\kappa^2(z)-k_n^2}$ is the vertical wave number, and A is a normalization constant. The field incident on the surface due to mode n from source element j is obtained by replacing $\Phi_n(z)$ in the expression for the nth mode in Eq. (B1) by $\Phi_n^+(0)$:

$$P_n^{\text{inc}}(r,z_j) = (2i\pi)^{1/2}\rho^{-1}\frac{1}{\sqrt{k_n r}}\Phi_n(z_j)\Phi_n^+(0)e^{(ik_n-\delta_n)r}, \tag{B3}$$

where z_j is the source element depth. The contribution of the nth mode to the incident field produced by a steered source array is then given by

$$P_n^{\text{inc}}(r,z_s) = \frac{1}{J}\sum_{j=1}^{J} P_n^{\text{inc}}(r,z_j)e^{i\beta_j}, \tag{B4}$$

where J is the number of source elements, z_s is now the effective array depth, and β_j is a plane wave element phasing. By a similar argument, the mth mode of the scattered field at the receiver depth z_r is

$$P_m^{\text{scatt}}(r,z_r) = (2i\pi)^{1/2}\rho^{-1}\frac{1}{\sqrt{k_m r}}\Phi_m(z_r)\Phi_m^-(0)e^{(ik_m-\delta_m)r}. \tag{B5}$$

The pressure at the receiver is expressed as a weighted sum over the scattered modes with weightings determined by the incident mode amplitudes and mode scattering matrix S_{nm}:

$$P_{\text{rec}}(r,\phi) = \sum_n\sum_m P_n^{\text{inc}}(r,z_s)S_{nm}P_m^{\text{scatt}}(r,z_r). \tag{B6}$$

The mode scattering matrix determines the contribution per unit area of the nth incident mode to the effective ''source strength'' of the mth scattered mode.

In general the mode scattering matrix is a complex function with a random phase and amplitude. The mean-square value of S_{nm} is equivalent to the plane wave scattering strength evaluated at the modal equivalent angles θ_n and θ_m, that is,

$$\langle |S_{nm}|^2\rangle = \sigma(\theta_n,\theta_m), \tag{B7}$$

where $\langle \cdot \rangle$ denotes an ensemble average, σ is the plane wave scattering strength, and $\theta_n = \cos^{-1}(c_n/c_0)$ with c_n being the modal phase speed and c_0 being the sound speed near the surface. Since the plane wave scattering strength σ is usually the available quantity, it is usually assumed that $S_{nm} = \sqrt{\langle |S_{nm}|^2 \rangle} = \sigma^{1/2}$. Using Eq. (B6) and Eq. (B7) gives

$$P_{\mathrm{rec}}(r,\phi) = \sum_n \sum_m P_n^{\mathrm{inc}}(r,z_s) \sqrt{\sigma(\theta_n,\theta_m;\phi)} P_m^{\mathrm{scatt}}(r,z_r). \tag{B8}$$

Equation (20) for the beam intensity $I(r,\phi)$ was derived assuming a monochromatic signal, whereas a moving surface imparts a frequency shift and a small frequency spread (less than 0.5 Hz in our case) due to the spectrum of water wave components traveling at different phase speeds. Because each of these frequency effects is small relative to the carrier frequency, Eq. (20) may be used with the down- or up-Bragg scattering strength to estimate the intensity in each of those narrow bands. Because the spreading parameter β is essentially constant over our range of Bragg frequency shifts, the approximation $h(\phi_D,\Omega) \approx h(\phi_D,\Omega_0)$ is permitted, where Ω_0 is a representative Bragg shift. Using Eqs. (15), (16), and (17) the integral for $I(r,\phi)$ can be re-expressed as a product of independent polar and range integrations, giving expressions for the down- and up-Bragg intensities:

$$I(r,\phi) = I_0 H(\phi) \int_{r-(c\tau/4)}^{r+(c\tau/4)} |P_{\mathrm{rec}}(r')|^2 r'\, dr', \tag{B9}$$

where

$$H(\phi) = \int_{-\pi}^{\pi} h(\phi',\Omega_0) B(\phi,\phi') d\phi' \quad \text{for down Bragg,} \tag{B10}$$

$$H(\phi) = \int_{-\pi}^{0} h(\phi'+\pi,\Omega_0) B(\phi,\phi') d\phi' + \int_{0}^{\pi} h(\phi'-\pi,\Omega_0) B(\phi,\phi') d\phi' \quad \text{for up Bragg,} \tag{B11}$$

and the range component of the received pressure is given by

$$P_{\mathrm{rec}}(r) = \sum_n \sum_m P_n^{\mathrm{inc}}(r,z_s) \sqrt{\sigma_v(\Omega_{nm})} P_m^{\mathrm{scatt}}(r,z_r), \tag{B12}$$

with

$$\Omega_{nm} = (g)^{1/2} (\kappa_s \cos\theta_n + \kappa_i \cos\theta_m)^{1/2}. \tag{B13}$$

The azimuthal integral [Eqs. (B10) and (B11)] is straightforward to compute numerically using the array beam pattern and the spreading function $h(\phi_D,\Omega_0)$. To avoid the singularities at the modal turning points of the up- and down-going waves [$\gamma_n = 0$ in Eq. (B2)], modes with incident angles below 2° are excluded from the sum. These modes may be omitted since backscattering strengths are extremely small at such low grazing angles.

[1] P. M. Ogden and F. T. Erskine, ''Surface scattering measurements using broadband explosive charges in the Critical Sea Test Experiments,'' J. Acoust. Soc. Am. **95**, 746–761 (1994).

[2] P. M. Ogden and F. T. Erskine, ''Surface and volume scattering measurements using broadband explosive charges in the Critical Sea Test 7 experiments,'' J. Acoust. Soc. Am. **96**, 2908–2919 (1994).

[3] R. C. Gauss, J. M. Fialkowski, and R. J. Soukup, ''Low-frequency direct-path surface and volume scattering measured using narrowband and broadband pulses,'' in *Ocean Reverberation*, edited by D. Ellis, J. Preston, and H. Urban (Kluwer Academic, Boston, 1993), pp. 235–240.

[4] R. C. Gauss, J. M. Fialkowski, and R. J. Soukup, ''Intermediate-Pulse Direct-Path measurements of Surface Scattering in CST-7 Phase II,'' in *Critical Sea Test 7, Phase II - Principal Investigators' Results, Volume 1: Surface Scattering and Volume Scattering Measurements*, edited by F. T. Erskine, STD-R-2258 (Johns Hopkins University, Applied Physics Laboratory, 1993), pp. 3-1 to 3-68.

[5] R. C. Gauss, P. M. Ogden, J. B. Chester, and J. M. Fialkowski, ''Deriving scattering strengths from nonstationary time-series data: A comparison of low-frequency surface-backscattering strengths using both implusive and coherent sources,'' J. Acoust. Soc. Am. **97**, 3403(A) (1995).

[6] M. A. Donelan, J. Hamilton, and W. H. Hui, ''Directional spectra of wind-generated waves,'' Philos. Trans. R. Soc. London, Ser. A **315**, 509–562 (1985).

[7] M. L. Banner, ''Equilibrium spectra of wind waves,'' J. Phys. Oceanogr. **20**, 966–984 (1990).

[8] D. R. Jackson, D. P. Winebrenner, and A. Ishimaru, ''Comparison of perturbation theories for rough-surface scattering,'' J. Acoust. Soc. Am. **83**, 961–969 (1988).

[9] E. I. Thorsos, ''Acoustic scattering from a ''Pierson–Moskowitz'' sea surface,'' J. Acoust. Soc. Am. **88**, 335–349 (1990).

[10] M. Pourkaviani and J. F. Willemsen, ''Perturbative corrections to acoustic Doppler backscattering I: Ocean surface gravity waves,'' J. Acoust. Soc. Am. **90**, 426–432 (1991).

[11] I. M. Fuks, ''Spectral width of signals scattered by a disturbed sea surface,'' Sov. Phys. Acoust. **20**, 276–281 (1974).

[12] J. T. Warfield, ''Doppler shifting of surface scattered reverberation,'' U. S. Navy Journal of Underwater Acoustics 183–205 (1981).

[13] J. L. Hanson and L. H. White, ''Assessment of the Air–Sea Boundary Zone During Critical Sea Test 4,'' STD-R-1978, JHU/APL Report (1991).

[14] E. I. Thorsos, ''The validity of the Kirchhoff approximation for rough surface scattering using a Gaussian roughness spectrum,'' J. Acoust. Soc. Am. **83**, 78–92 (1988).

[15] N. Bleistein and R. A. Handelsman, *Asymptotic Expansions of Integrals* (Dover, New York, 1986), pp. 342–347.

[16] R. J. Urick, *Principles of Underwater Sound* (McGraw-Hill, New York, 1983).

Influence of perturbations on chaotic behavior of the parabolic ray system

Xiaojun Li, Yu Zhang, and Gonghuan Du
Institute of Acoustics, State Key Laboratory of Modern Acoustics, Nanjing University, Nanjing, 210093, People's Republic of China

(Received 18 December 1997; revised 19 July 1998; accepted 15 December 1998)

This paper points out that the realistic sound speed is not only a periodic function of range; there are also many other factors such as random perturbation modulating the periodic perturbation which had been investigated in previous papers. By the calculations of phase portraits, power spectra, Lyapunov exponents, etc., the influence of random perturbation on the dynamic behavior of the ray system is investigated. The study demonstrates that the random perturbation lowers the threshold from quasiperiod to chaos, and reinforces the irregularity of the trajectories which were regular originally. However, for the trajectories which were chaotic originally, the random perturbation makes their irregularity decrease slightly. Finally, the probability distribution of the Lyapunov exponent is given. © *1999 Acoustical Society of America.* [S0001-4966(99)06703-X]

PACS numbers: 43.30.Cq, 43.25.Rq [SAC-B]

INTRODUCTION

Chaotic or turbulent behavior can be seen in many systems which are governed by deterministic equations in fields such as physics, chemistry, and biology. Recently, in underwater acoustics, it has been pointed out that chaotic behavior can occur in a class of ray systems.[1–7] Experiments have been performed which show propagation behavior consistent with chaotic ray behavior.[8–11] When the sound speed is independent of range, the integrability of a ray system guarantees regular ray motion. However, when the influence of some perturbations caused by factors such as internal waves, mesoscale structure, etc. is considered, the range-dependent sound-speed model may lead to chaotic behavior.[1,3] As is known, the important characteristic of a chaotic dynamic system is its sensitivity to the perturbation; it exhibits a continuous instability with the perturbation of the parameters, particularly at the points close to the threshold of instability. Any infinitesimal perturbation may lead to remarkable deviation of the trajectories. Naturally, the influence of some perturbation on the ray system caused by factors such as internal waves, mesoscale structure, etc. should be investigated. Furthermore, random perturbation exists in the real physical systems everywhere. Any experiment process must be affected by at least two factors: fluctuation of parameters and external random perturbation. Recent work has demonstrated that the existence of random perturbation may lead to rich dynamic behavior of the nonlinear system; for example, it can make the trajectory jump between the different basins of attraction,[12] lower the threshold value of the parameter for the onset of chaos,[13] and significantly amplify the small resonant periodic perturbation in the presence of noise. This is known as stochastic resonance.[14] Thus, in the acoustic ray model, the effects that random fluctuation forces have on both the nature of the chaotic attractor and its asymptotic state should be investigated. In this paper, we investigate the dynamic behavior under the influence of random perturbation with the help of various tools such as phase portraits, power spectra, Lyapunov exponents, etc. In the numerical calculation, the reflection at the ocean surface is investigated with a method different from what is suggested by Smith *et al.*[1] We find that, for the trajectories which are regular originally, the random perturbation reinforces their irregularity and leads them from quasiperiod to chaos, but for those trajectories which are irregular originally, the random perturbation makes their irregularity decrease slightly. Based on the calculation of the probability distribution of Lyapunov exponents over the ensemble of random fluctuation forces, we investigate the moments of Lyapunov exponents, which are more effective to determine ray chaos under random perturbation.

I. SYSTEM MODEL

High-frequency approximation makes it possible for underwater sound waves to be described by a ray system. Furthermore, according to the ''standard PE'' approximation, under the condition of small ray angle with respect to the horizontal, the ray system can be written in a parabolic form[2,15]

$$\frac{dz}{dr}=\frac{\partial H}{\partial p}, \tag{1a}$$

$$\frac{dp}{dr}=-\frac{\partial H}{\partial z}, \tag{1b}$$

where z and r are depth and range, respectively, $p=\tan\theta$, θ is the ray angle with respect to the horizontal, and

$$H(z,p,r)=p^2/2+[c(z,r)-c_0]/c_0 \tag{2}$$

is the Hamiltonian; here c_0 is the reference sound speed.

The sound speed is a fundamental quantity in underwater acoustics and has a very important effect on ray behavior. Considering some perturbation factors to sound speed, we write the sound-speed model as

$$c(z,r)=f(z)+\varepsilon g(z,r)+\delta w(z,r), \tag{3}$$

 0001-4966/99/105(4)/2142/7/$15.00 © 1999 Acoustical Society of America

where $f(z)$ is the background range-independent sound-speed profile, $g(z,r)$ and $w(z,r)$ are the perturbations, and ε and δ are the perturbation coefficients. For $f(z)$, in the following, we restrict our discussion to the double-channel sound-speed profile which is fit for data from measurement made on the North Atlantic double-channel system[2,16]

$$f(z)=a_0+a_1z+a_2z^2+a_3z^3+a_4z^4, \tag{4}$$

where a_0, a_1, a_2, a_3, and a_4 are all constants, and $a_4>0$. The first perturbation $g(z,r)$, following the discussion of Palmer *et al.*,[17] is periodic range-dependent

$$g(z,r)=\sqrt{2}\exp(-1.5z/B)\sin(2\pi r/R), \tag{5}$$

where R is the perturbation period, and B is the depth scale. This perturbation can be interpreted as a highly idealized internal wave or a single baroclinic mode representing mesoscale structure.[2,3,17] The additional perturbation $w(z,r)$ is caused by other reasons.

Substituting Eqs. (2), (3), (4), and (5) into Eq. (1) gives

$$\frac{dz}{dr}=p, \tag{6a}$$

$$\frac{dp}{dr}=-\frac{1}{c_0}\left[a_1+2a_2z+3a_3z^2+4a_4z^3 -\varepsilon\sqrt{2}\left(\frac{1.5}{B}\right)\exp\left(\frac{-1.5z}{B}\right)\sin\left(\frac{2\pi r}{R}\right)+\delta\frac{\partial w(z,r)}{\partial z}\right]. \tag{6b}$$

It can be noted that, if the last term at the right-hand side of Eq. (6b) is ignored, Eq. (6) is the same model as what was introduced in Ref. 2.

Previous research has studied the case that only the periodic range-dependent perturbation exists,[1,2] i.e., $c(z,r)=f(z)+\varepsilon g(z,r)$. Their results demonstrated the route from quasiperiod to chaos. Here, more details are presented about this route to chaos, which is different from the period-doubling route of Feigenbaum[18] and the intermittent route of Pomean–Manneville.[19] Consider the following unperturbed system equation, which is obtained by setting all of the perturbations to zero in Eq. (3)

$$\frac{d^2z}{dr^2}=-\frac{1}{c_0}(a_1+2a_2z+3a_3z^2+4a_4z^3). \tag{7}$$

Equation (7) is a second-order linear ordinary differential equation with constant coefficients, and its solution has a natural frequency ω_0. When the initial ray angle θ_0 is not large, e.g., $\theta_0=2$ deg, nonlinearity is weak so that the trajectory is attracted within its own basin of attraction; it is periodic. When $g(z,r)$, which is a periodic function with a frequency $\omega=2\pi/R$ is added, mostly, because the ratio of ω_0 to ω is an irrational number, the trajectory wanders within a two-dimension torus and gradually fills with it. This trajectory is quasiperiodic. At this time, if the ray system is perturbed by some other additional perturbations, the quasiperiodic trajectory may become unstable so that the invariant torus breaks up; chaotic behavior then appears.

The ocean environment is nonuniform and time-varying. There are a great number of blocks of water with certain temperature and scale which have random distributions inside the ocean. Other accidental factors may also have effects on the ray behavior. All of these factors make the sound speed in the ocean have random characteristics. We take the aforementioned additional perturbation as a random term with Gaussian distribution. Replacing the last term at the right-hand side of Eq. (6b) by a stationary random process with Gaussian distribution gives equations as below.

$$\frac{dz}{dr}=p, \tag{8a}$$

$$\frac{dp}{dr}=-\frac{1}{c_0}\left[a_1+2a_2z+3a_3z^2+4a_4z^3 -\varepsilon\sqrt{2}\left(\frac{1.5}{B}\right)\exp\left(\frac{-1.5z}{B}\right)\sin\left(\frac{2\pi r}{R}\right)\right]+\xi(r), \tag{8b}$$

where $\xi(r)$ is a stochastic process, which is a Gaussian random variable with mean value 0 and mean squared value σ^2 in a certain value of range r.

Our study demonstrates that when random perturbation with certain strength is added to the system with quasiperiodic trajectories, these original regular trajectories may exhibit chaotic behavior. So, we can deduce that the random perturbation lowers the threshold of chaos. But how does the system evolve from quasiperiod to chaos? It is known that there are three fixed points in the phase plane of the ray system used in this paper;[3] the one in the middle is a saddle point, while the upper and the lower ones are center points. When the system is unperturbed, the above-mentioned trajectories are quasiperiodic and move within their own basins of attraction in a regular fashion. After the random perturbation is added, the trajectories begin to go away from their original orbits. When the perturbation becomes strong enough, trajectories may accidentally escape out of the basin of attraction. It approaches the saddle point, but the saddle point is unstable, and small deviations or perturbation will cause the trajectory to move away from it. The saddle point repels the trajectory to the basin of attraction from which it comes or to the other basin of attraction at random. An opposite process may occur in the meantime. So, the trajectories jump up and down between the two basins of attraction in a random-like fashion. In other words, random perturbation reinforces the overlap of resonance structure, and the conditions of the KAM theorem begin to break down, which leads to the appearance of stochastic behavior, i.e., chaos appears.

One must note that, when the ray angle is not very small, the ray may intersect the ocean surface. In order to describe behavior of the trajectories correctly, reflection of the trajectory at the ocean surface should be considered. In previous papers,[1] the intersection of the trajectory and the ocean surface was found iteratively. This requires quite a few times of calculation to approximately reach the intersection. Here, we use the method suggested by Henon,[20] which can obtain the intersection with only one single step of calculation; particularly, this method can obtain the intersection exactly while the iterative method cannot. When the iteration comes near the surface $(z=0)$ within a step, i.e., if we find at the nth

step r_n, p_n, $z_n>0$, and the next step $r_n+\Delta r$, p_{n+1}, $z_{n+1}<0$, the intersection must occur between these two steps. Interchanging the role of z and r, then dividing Eq. (8b), by Eq. (8a), and inverting Eq. (8a), we obtain the following equation:

$$\frac{dr}{dz}=\frac{1}{p}, \tag{9a}$$

$$\frac{dp}{dz}=-\frac{1}{c_0 p}\left[a_1+2a_2z+3a_3z^2+4a_4z^3 - \varepsilon\sqrt{2}\left(\frac{1.5}{B}\right)\exp\left(\frac{-1.5z}{B}\right)\sin\left(\frac{2\pi r}{R}\right)\right]+\frac{\xi(r)}{p}. \tag{9b}$$

Integrating Eq. (9) in the new independent variable z by one step $\Delta z=-z_n$ from the point z_n using the initial value r_n, p_n, one reaches exactly the ocean surface $z=0$ plane. After changing the sign of p, we restore the role of variables, and continue integrating normally. Here, we actually consider that the ocean surface is flat and has a mirror reflection.

II. NUMERICAL CALCULATIONS

For system (8), a variety of numerical experiments such as phase portraits ($p-z$ curve), time history ($r-z$ curve), power spectra, and Lyapunov exponents are performed. In numerical studies, a fourth-order Runge–Kutta method is applied. In our calculation, value of the coefficients and parameters used are

$$a_0=1.493\,23,\quad a_1=-0.047,\quad a_2=0.147\,14,$$

$$a_3=-0.145\,17,\quad a_4=0.045\,11,$$

according to data taken from the North Atlantic double-channel system, and

$$c_0=1.5\ \text{km/s},\quad B=1.0\ \text{km},\quad \varepsilon=0.0025,$$

$$R=50\ \text{km}.$$

The sound source is placed near the lower channel axis of the double-channel model, so initial depth $z_0=1.345$ km. Initial ray angle is $\theta_0=2$ deg or $\theta_0=5$ deg. It should be mentioned that even if the parameters and the strength of perturbation do not change, different initial conditions can also lead to different behavior of the ray system.

Shown in Fig. 1 is the time history under different initial ray angles and strength of random perturbations. In parts (a) and (c), there is no perturbation, i.e., $\sigma=0$. In parts (b) and (d), the perturbation has a strength $\sigma=0.01$. We can see that, when $\theta_0=2$ deg [parts (a) and (b)], for the unperturbed system, the trajectory is regular; however, in the presence of perturbation, the trajectory becomes irregular. Then, with the increasing of the perturbation, the irregularity of the trajectory will continue to be reinforced. When $\theta_0=5$ deg [parts (c) and (d)], the system has already been chaotic before the perturbation is added; with increasing perturbation, the irregularity decreases slightly.

Phase portraits are given in Fig. 2. Arrangement of the parts of Fig. 2 is similar to that of Fig. 1. In part (a) ($\theta_0=2$ deg, $\sigma=0$), we find that the phase portrait is a regular trajectory around the center of the lower basin of attraction,

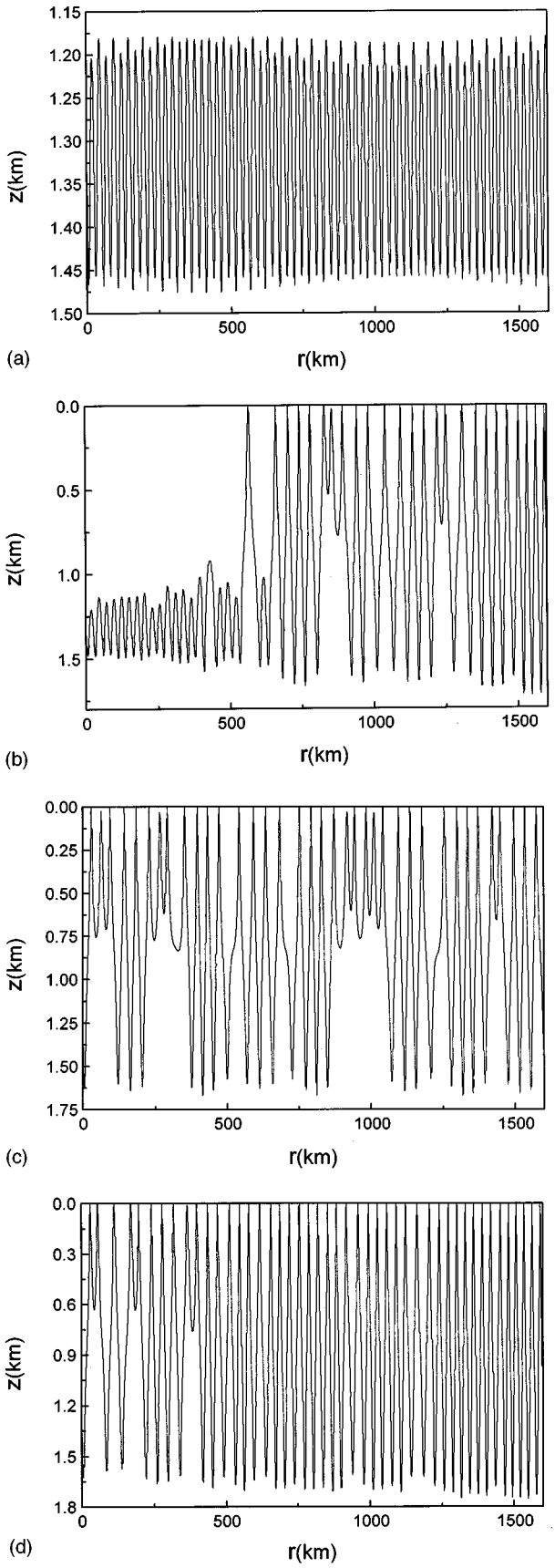


FIG. 1. The time history, (a) $\theta_0=2$ deg, $\sigma=0$, regular trajectory, (b) $\theta_0=2$ deg, $\sigma=0.01$, irregular trajectory under random perturbation, (c) $\theta_0=5$ deg, $\sigma=0$, irregular trajectory, (d) $\theta_0=5$ deg, $\sigma=0.01$, irregular trajectory under random perturbation.

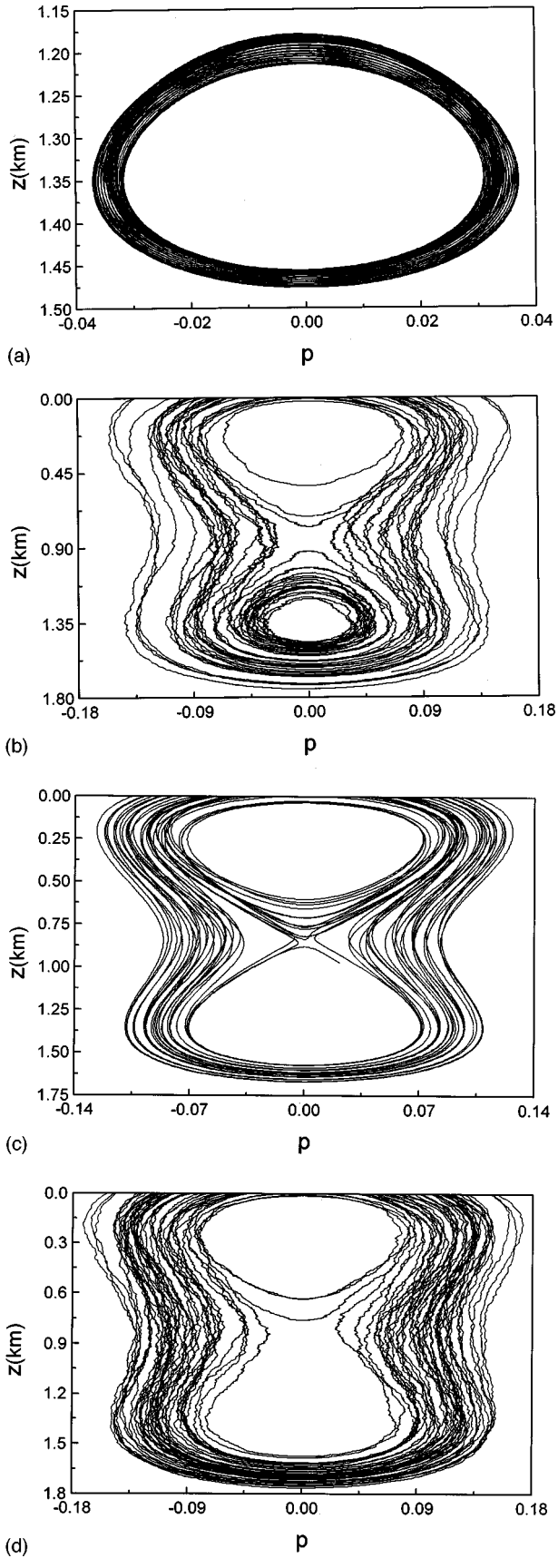


FIG. 2. The phase portraits, (a) $\theta_0=2$ deg, $\sigma=0$, quasiperiodic trajectory, (b) $\theta_0=2$ deg, $\sigma=0.01$, chaotic trajectory under random perturbation, (c) $\theta_0=5$ deg, $\sigma=0$, chaotic trajectory, (d) $\theta_0=5$ deg, $\sigma=0.01$, chaotic trajectory under random perturbation.

and it represents quasiperiod. With the perturbation added, the trajectory has a fluctuation on the previous quasiperiodic path. When random perturbation become larger, the trajectory begins to lose its way and jumps between the two basins of attraction; the system shows chaotic behavior. When $\theta_0=5$ deg, the system is chaotic at the beginning, and the random perturbation makes the irregularity of the system decrease slightly. Here we note that, because of reflection, p changes its sign at $z=0$, so that the trajectories in phase portrait have some interval at the surface. In the time history, we can see the mirror-reflection process clearly.

Another tool that we use to describe the behavior of nonlinear systems is the power spectrum. A power spectrum shows the amplitude of all harmonic components of solution of the system; it also can tell chaos from regularity. Power spectra are shown in Fig. 3. Under the initial ray angle $\theta_0=2$ deg [parts (a) and (b)], we can see that when the system is unperturbed, the power spectrum is almost discrete. So, we can deduce that the trajectories are regular. While there is a random perturbation with certain strength, it shows broadband fashion. This indicates that the system motion is irregular. Furthermore, when the strength of the random perturbation increases, the band of the power spectrum becomes wider and wider. In the case of $\theta_0=5$ deg [parts (a) and (b)], with the random perturbation varying, the power spectrum remains the broadband fashion.

As known, there have been many approaches to quantify chaos, e.g., phase portrait, power spectrum, entropy, and fractal dimension; however, the spectra of Lyapunov exponents was proven to be the most useful dynamic diagnostic for chaotic systems. Lyapunov exponents are relative to the average exponential rates of divergence or convergence of neighboring orbits in phase space. Any system containing at least one positive Lyapunov exponent is defined to be chaotic, while the system with no positive exponent is regular. From the above calculations of the phase portraits and the power spectra, we find that the dynamic behavior of the ray system is closely related to the random perturbation. Thus, the influence of random perturbation on the ray model should be investigated by the calculation of Lyapunov exponents.

Here, we extend the previous discussion about the calculation of the Lyapunov exponents in Ref. 21 to the nonlinear dynamic system with the additive noise. The dynamic system can be described as

$$\frac{d\mathbf{X}}{dt}=F(\mathbf{X})+\boldsymbol{\eta}, \tag{10}$$

where $\mathbf{X}$ is an n-dimension vector in the state space, $\boldsymbol{\eta}$ is an n-dimension random vector, and t is the time. In the ray system which is presented in this paper, $\mathbf{X}=(z,p)^T$, $\boldsymbol{\eta}=[0,\xi(t)]^T$, t is corresponding to the range r, and $\xi(t)$ is a random variable with mean value 0 and mean squared value σ^2. On the other hand, the evolution of the tangent vector $\mathbf{W}$ in the tangent space of $\mathbf{X}(t)$ is given by the linearization of Eq. (10),

$$\delta\mathbf{W}=F_x(\mathbf{W})\,\delta\mathbf{W}, \tag{11}$$

where $\mathbf{W}$ is the n-dimension eigenvector which represents the evolution of a coordinate system corresponding to the

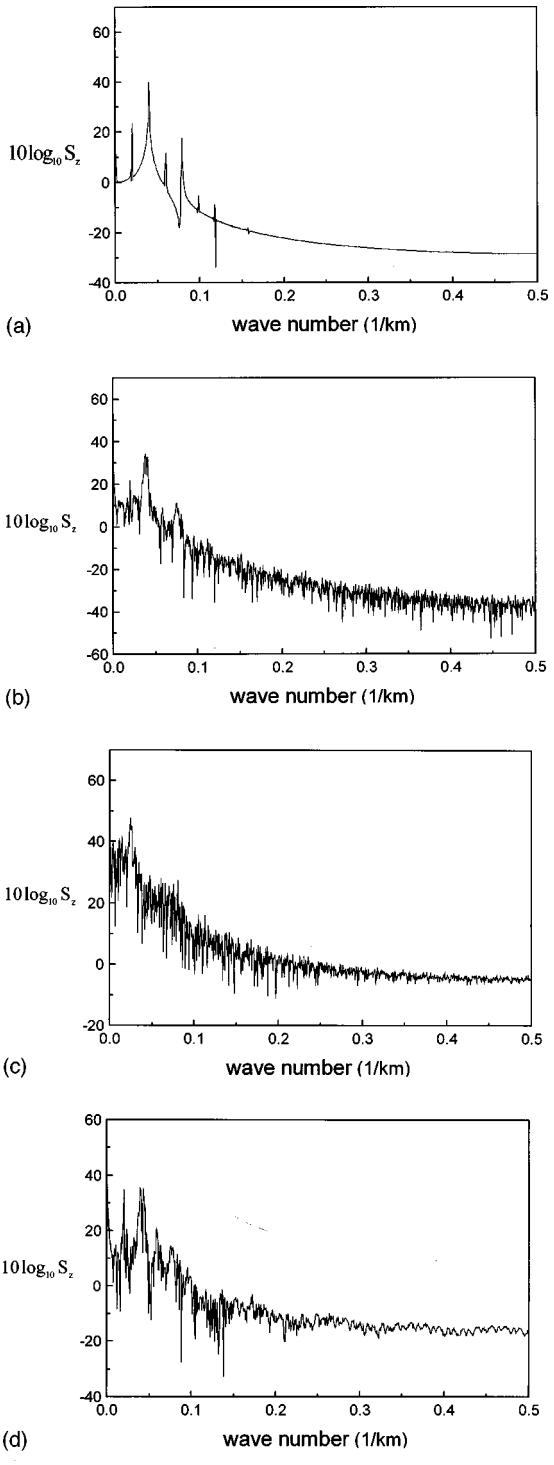


FIG. 3. The power spectra, where S_z is the power of the depth z, (a) $\theta_0=2$ deg, $\sigma=0$, regular trajectory, (b) $\theta_0=2$ deg, $\sigma=0.01$, chaotic trajectory under random perturbation, (c) $\theta_0=5$ deg, $\sigma=0$, chaotic trajectory, (d) $\theta_0=5$ deg, $\sigma=0.01$, chaotic trajectory under random perturbation.

Lyapunov exponent, and $\delta\mathbf{W}$ is the significantly small disturbance to $\mathbf{W}$. Let $\delta\mathbf{W}$ be the n-dimensional basis of the tangent space; then, the Lyapunov exponent λ_i are given by

$$\lambda_i=\lim_{t\to\infty}\frac{1}{t}\ln\frac{\|\delta\mathbf{W}_i(t)\|}{\|\delta\mathbf{W}_i(t_0)\|}\quad(i=1,2,...,n),\tag{12}$$

where $\mathbf{W}_i(t)$ is the ith basis vector in tangent space, and t_0 is the initial time. Because the evolution of phase space is related to the random fluctuation, the Lyapunov exponent has some statistical characteristic for the small Gaussian random perturbation. Generally, it is difficult to obtain an exact statistical model of the Lyapunov exponent directly. Kapitaniak *et al.*[22] had suggested a method to calculate the random Lyapunov exponents by approximating the stochastic process $\boldsymbol{\eta}$ as a sum of harmonic components. Their calculation demonstrated that the Lyapunov exponents have a certain probability distribution. However, their expanding the random function as the sum of harmonic components is obviously not strict. Here, we would like to calculate the distribution of Lyapunov exponents over the ensemble of random fluctuation forces $\xi(t)$. We regard the Lyapunov exponents as random variables. According to the moment theorem, if all the moments of a random variable are known, its distribution density is determined uniquely.[23] We obtain all the moments of Lyapunov exponent as follows:

$$\text{first-order moment}\quad E(\lambda_i)=\int_{-\infty}^{\infty}\lambda_i f(\zeta)d\zeta,\tag{13a}$$

$$\text{second-order moment}\quad D(\lambda_i)=\int_{-\infty}^{\infty}[\lambda_i-E(\lambda_i)]^2 f(\zeta)d\zeta,\tag{13b}$$

kth-order moment

$$E[\lambda_i-E(\lambda_i)]^k=\int_{-\infty}^{\infty}[\lambda_i-E(\lambda_i)]^k f(\zeta)d\zeta,\tag{13c}$$

where $f(\zeta)$ is the probability density of the Lyapunov exponent.

But for a chaotic system, because of the stretched, folded characteristics, $F_{\mathbf{W}}(\mathbf{X})$ is time-dependent so that it is very difficult to solve Eq. (13) with the analytical method. In general, we have to feed back on the numerical method, and use the discrete scheme approaching the continuous distribution of the Lyapunov exponents. Calculating Lyaponov exponents with respect to different random forces, we divide the region between the lower-boundary λ_i^{sub} and the upper-boundary λ_i^{sup} of the ith Lyapunov exponent λ_i into m equal subregions. Length of the subregion is

$$\Delta\lambda_i=\frac{\lambda_i^{\text{sup}}-\lambda_i^{\text{sub}}}{m}.\tag{14}$$

$\lambda_i(j)$ $(j=0,1,...,m-1)$ is defined as $\lambda_i(j)=[j\Delta\lambda_i+(j+1)\Delta\lambda_i]/2$ in the jth subregion, i.e., $j\Delta\lambda_i\sim(j+1)\Delta\lambda_i$. The probability that λ_i appears between $j\Delta\lambda_i$ and $(j+1)\Delta\lambda_i$ is

$$P[\lambda_i(j)]=\frac{\text{num}(j)}{n_t},\tag{15}$$

where $\text{num}(j)$ is the number of times that λ_i appears in the jth subregion, which is obtained from numerical calculation, and n_t is the total number of λ_i. According to this, we can obtain the distribution $P(\lambda_i)\sim\lambda_i$ curve. Then, the moments of the Lyapunov exponents can be calculated as

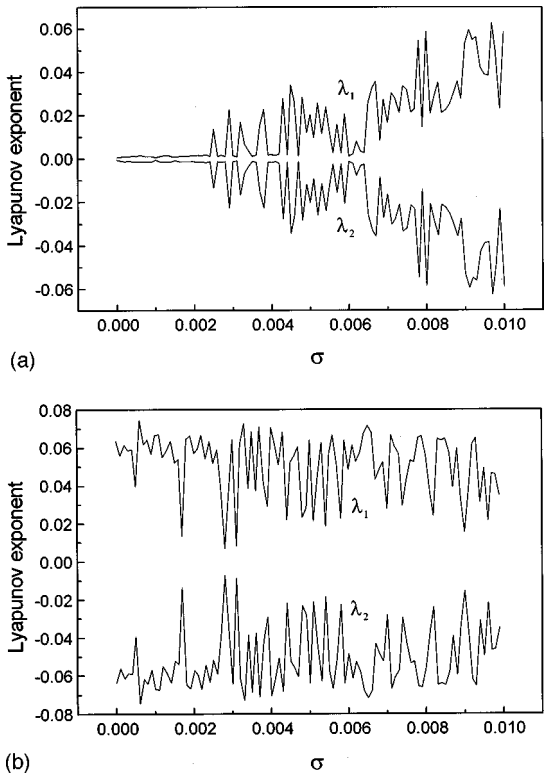


FIG. 4. Lyapunov exponents, where λ_1 is the maximum Lyapunov exponent, λ_2 is the second Lyapunov exponent, (a) $\theta_0=2$ deg, the trajectory evolves from quasiperiod to chaos, (b) $\theta_0=5$ deg, irregularity of the trajectory decreases slowly.

$$E(\lambda_i)=\sum_{j=0}^{m-1}\lambda_i(j)P[\lambda_i(j)], \tag{16a}$$

$$D(\lambda_i)=\sum_{j=0}^{m-1}[\lambda_i(j)-E(\lambda_i)]^2P[\lambda_i(j)], \tag{16b}$$

$$E[\lambda_i-E(\lambda_i)]^k=\sum_{j=0}^{m-1}[\lambda_i(j)-E(\lambda_i)]^kP[\lambda_i(j)]. \tag{16c}$$

From above, we perform the calculation of Lyapunov exponents in the ray system. Figure 4 shows the Lyapunov exponents on the perturbation strength σ. The initial ray angle $\theta_0=2$ deg and $\theta_0=5$ deg versus part (a) and part (b), respectively. Here, λ_1 is the maximum Lyapunov exponent and λ_2 is the second Lyapunov exponent. It is clear that the regular trajectory with $\sigma=0$ will be led to chaos by random perturbation with the increase of σ, while the chaotic attractor is stable under the influence of small fluctuation. This shows agreement with the above calculation. Furthermore, in the case of random excitation, the properties of the Lyapunov exponents are obtained from the realization of the random process $\eta(t,\omega)$, and their distributions allow one to quantify chaos. Here we obtain two types of distribution. The first type of Lyapunov exponents, whose distribution value is mainly zero, corresponds to the regular behavior of the system, while the second one, which is shown in Fig. 5, corresponds to the chaotic behavior. In Fig. 5, the initial ray angle $\theta_0=5$ deg, the random perturbation strength $\sigma=0.01$, the solid and hollow points represent the numerical results of

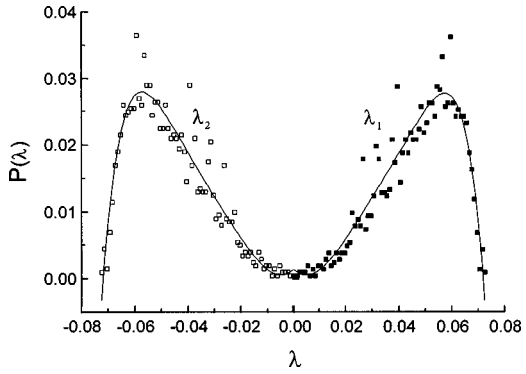


FIG. 5. The probability distribution of Lyapunov exponents ($\theta_0=5$ deg, $\sigma=0.01$), where $P(\lambda)$ represents the probability of the Lyapunov exponents λ, the solid and hollow points represent the numerical results of Lyapunov exponents λ_1 and λ_2, respectively, and the solid lines are their sixth-order polynomial fitting.

Lyapunov exponents λ_1 and λ_2, respectively, and the solid lines are their sixth-order polynomial fitting. The mean value of the maximum Lyapunov exponent λ_1 is $E(\lambda_1)=0.0474$, and its mean square value is $D(\lambda_1)=0.0002$.

III. SUMMARY

We have investigated the ray-system behavior influenced by random perturbation with several numerical tools such as phase portraits, power spectra, Lyapunov exponents, etc. Through this work, we found that with adding and increasing the random perturbation, the trajectories which were previously regular became irregular; however, irregularity of the trajectories which were previously irregular decreased slightly. The probability distribution of Lyapunov exponents was calculated. All of these studies provided references to describe the ray chaos, predict the path of ray propagation, and measure the sound signal in deep oceans over long range.

[1] K. B. Smith, M. G. Brown, and F. D. Tappert, ''Ray chaos in underwater acoustics,'' J. Acoust. Soc. Am. **91**, 1939–1949 (1992).
[2] Z. Jiang, T. A. Pitts, and J. F. Greenleaf, ''Analytic investigation of chaos in a class of parabolic ray systems,'' J. Acoust. Soc. Am. **101**, 1971–1980 (1997).
[3] K. B. Smith, M. G. Brown, and F. D. Tappert, ''Acoustic ray chaos induced by mesoscale ocean structure,'' J. Acoust. Soc. Am. **91**, 1939–1949 (1992).
[4] F. D. Tappert, M. G. Brown, and G. Goñi, ''Weak chaos in an area-preserving mapping for sound ray propagation,'' Phys. Lett. A **153**, 181–185 (1991).
[5] M. G. Brown, F. D. Tappert, and G. Goñi, ''An investigation of sound ray dynamics in the ocean volume using an area preserving mapping,'' Wave Motion **14**, 93–99 (1991).
[6] M. A. Wolfson and F. D. Tappert, ''Chaos in an acoustic propagation model,'' in *The Chaos Paradigm: Developments and Applications in Engineering and Science*, edited by R. A. Katz (AIP Conf. Proc. 296, AIP, New York, 1994), pp. 277–288.
[7] D. Tappert and X. Tang, ''Ray chaos and eigenrays,'' J. Acoust. Soc. Am. **99**, 185–195 (1996).
[8] P. Worcester, B. Dushaw, and B. Howe, ''Gyre-scale reciprocal acoustic transmissions,'' in *Ocean Variability and Acoustic Propagation*, edited by J. Potter and A. Warn-Varnas (Kluwer Academic, Dordrecht, 1990).
[9] J. L. Spiesberger and K. Metzger, ''A new algorithm for sound speed in water,'' J. Acoust. Soc. Am. **89**, 2677–2687 (1991).
[10] B. M. Howe, J. A. Mercer, R. C. Spindel, P. F. Worcester, J. A. Hildebrand, W. S. Hodgkiss, T. F. Duda, and S. M. Flatté, ''SLICE 89: A single

slice tomography experiment,'' in *Ocean Variability and Acoustic Propagation*, edited by J. Potter and A. Warn-Varnas (Kluwer Academic, Dordrecht, 1990).
[11] K. D. Heaney, W. A. Kuperman, and B. E. McDonald, ''Perth–Bermuda sound propagation (1960): Adiabatic mode interpretation,'' J. Acoust. Soc. Am. **90**, 2586–2594 (1991).
[12] F. T. Arecchi, R. Badii, and A. Politi, ''Generalized multistability and noise-induced jumps in a nonlinear dynamical system,'' Phys. Rev. A **32**, 402–408 (1985).
[13] J. P. Crutchfield and B. A. Huberman, ''Fluctuations and the onset of chaos,'' Phys. Lett. **77A**, 407–410 (1980).
[14] L. Gammaitoni, F. Marchesoni, E. Menichella-Saetta, and S. Santucci, ''Stochastic resonance in bistable system,'' Phys. Rev. Lett. **62**, 349–352 (1989).
[15] F. D. Tappert, ''The parabolic approximation method,'' in *Lecture Notes in Physics, Vol. 70, Wave Propagation and Underwater Acoustics*, edited by J. B. Keller and J. S. Papadakis (Springer, New York, 1977), pp. 224–287.
[16] C. A. Boyles, *Acoustic Waveguides, Application to Oceanic Science* (Wiley, New York, 1984), pp. 263–269.
[17] D. R. Palmer, M. G. Brown, F. D. Tappert, and H. F. Bezdek, ''Chaos in nonseparable wave propagation problems,'' Geophys. Res. Lett. **15**, 569–572 (1988).
[18] M. J. Feigenbaum, ''Quantitative universality for a class of nonlinear transformations,'' J. Stat. Phys. **19**, 25–52 (1978).
[19] Y. Pomean and P. Manneville, ''Intermittent transition to turbulence in dissipative dynamical systems,'' Commun. Math. Phys. **74**, 189–197 (1980).
[20] M. Henon, ''On the numerical computation of Poincare maps,'' Physica D **5**, 412 (1982).
[21] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, ''Determining Lyapunov exponents from a time series,'' Physica D **16**, 285–317 (1985).
[22] T. Kapitaniak, ''Chaos in a noisy mechanical system with stress relaxation,'' J. Sound Vib. **123**, 391–396 (1988).
[23] A. Papoulis, *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York, 1984).

Corrections to Foldy's effective medium theory for propagation in bubble clouds and other collections of very small scatterers

Frank S. Henyey[a)]
Applied Physics Laboratory, College of Ocean and Fishery Sciences, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

(Received 4 December 1996; revised 28 September 1998; accepted 10 December 1998)

Propagation through underwater bubble clouds is often discussed in terms of Foldy's effective medium theory of wave propagation through a random collection of a large number of point scatterers. This paper is concerned with corrections to that theory. This work was motivated by a paper of Ye and Ding, who provided the next term after Foldy's in an expansion in powers of Foldy's term for the effective wave number squared. When Foldy's term is not small, truncation after Ye and Ding's term overestimates the correction. A correction is derived which includes Ye and Ding's, but is much smaller than theirs when Foldy's term is large. The correction amounts to replacing the scattering amplitude in Foldy's expression with an effective scattering amplitude that differs from the original scattering amplitude by having the scattering loss calculated in the effective medium. © *1999 Acoustical Society of America.* [S0001-4966(99)05003-1]

PACS numbers: 43.30.Es, 43.20.Bi, 03.40.Kf [SAC-B]

INTRODUCTION

Waves propagate through a random collection of scatterers in a way similar to propagation through a medium with an index of refraction given by the density of scatterers and their strength of scattering. Fifty years ago, Foldy[1] presented an effective medium theory. Making plausible assumptions, he established

$$K^2 = k^2 + 4\pi N f, \tag{1}$$

where K is the wave number in the effective medium, k is what the wave number would be without the scatterers, N is the density of scatterers, and f is the scattering amplitude of a single scatterer. Foldy's model consists of a random (Poisson process) distribution of scatterers, each of which scatters isotropically (monopole, or s wave). This model is particularly appropriate for the propagation of sound in an underwater bubble cloud.

Investigations of bubble cloud acoustics have led to recent reexaminations of Foldy's theory. In this paper, a correction with a simple physical interpretation is derived. For applications to bubble clouds, this expression is small, and Foldy's original expression is probably adequate for any practical case. Other authors have proposed larger corrections.

In particular, Ye and Ding[2] have used the multiple scattering series, and proposed a correction to Foldy's expression. Their work provides the starting point for this paper. Ye and Ding calculated the next-order term in a double expansion in N and f to Eq. (1). The present work is based on the observation that $4\pi Nf/k^2$ is often large in interesting cases. For example, a cloud of underwater bubbles at frequencies well below resonance have $4\pi Nf/k^2$ on the order of 2×10^4 times the air volume fraction; many bubble clouds (see, e.g., Terrill and Melville[3]) have the air volume fraction larger than 10^{-4}. If $4\pi Nf/k^2$ is large, truncation of the double expansion is not justified. In particular, Ye and Ding's result requires that the propagation can be assumed to have the form e^{ikr}/r out to a distance of order $1/k$. If $K \gg k$, there is significant cancellation from the oscillations over this region, and the correction to Foldy's result is considerably smaller than that given by Ye and Ding's expression, by an amount of order k/K.

I. MULTIPLE SCATTERING SERIES

The calculations reported in the present paper are based on the multiple scattering expansion. Infinite classes of multiple scattering contributions are summed. Proper ''bookkeeping'' is done, so that no contribution is counted more than once, and so it is immediately evident whether or not a particular contribution has been included. These calculations are most conveniently carried out by a scattering diagram method. A simple diagram method was used by Ye and Ding.

The calculations presented here are (1) the multiple scattering contributions identified by Ye and Ding as giving Foldy's result, (2) Ye and Ding's result, and (3) the set of contributions that give the present result. Afterwards, simple additional contributions are discussed, although no formula is presented for the small correction to Foldy's result. As such, the results presented here are restricted to $N|f|^3 \ll 1$.

Before starting, the simplifying assumptions are listed. Foldy discussed both the mean field and the mean square fluctuations from that mean; the present paper is only concerned with the mean field. Foldy allowed arbitrary distributions of scatterer density; corrections to his expression can be simply expressed only for a uniform density, which is assumed here. For notational convenience, it is assumed that all scatterers have the same scattering amplitude f; the gen-

a)Electronic mail: frank@apl.washington.edu

eralization to a distribution of f values is straightforward, by use of the probability density of scattering amplitudes.

Because of the assumption of a uniform density of scatterers, the mean field from an isotropically radiating source must be a function of the radius r from the source, and not the direction. Any spatially homogeneous problem (away from the source which destroys the symmetry) with only one type of wave has solutions decomposable into ''irreducible'' solutions $e^{\mathbf{iK}\cdot\mathbf{x}}$, where the magnitude of $\mathbf{K}$ is independent of direction. All such solutions, and therefore their superposition, obey the equation $(\nabla^2+K^2)\Psi=0$. The only function of r that solves this equation away from the source and has outgoing waves at infinity is (up to a constant multiplier) e^{iKr}/r. Here K is referred to as the effective wave number, and turns out to be complex. The problem is to evaluate K in terms of the original wave number k, the density of scatterers N, and the scattering amplitude f.

Foldy showed that his model can be mathematically formulated in terms of an $n\times n$ matrix whose size n is the total number of scatterers. The field caused by any scatterer is $f(e^{ikr}/r)\Psi'$, where r is the distance from that scatterer and Ψ' is the total field at that scatterer with the exception of its own scattered field. Here Ψ' contains both the incident field and the field scattered by all the rest of the scatterers, each of which involves its own Ψ'. Thus, a set of linear equations can be written down which relate the vector of Ψ''s to the incident field. The matrix is made of the coefficients of this linear system of equations. The diagonal terms are unity, and the off-diagonal terms are $R_{\alpha\beta}=-f(e^{ikr}/r)$, where for the α, β element of the matrix, r is the distance between the αth and βth scatterers. The matrix has the form $\mathbf{I}-\mathbf{R}$, where $\mathbf{I}$ is the identity matrix and $-\mathbf{R}$ is the off-diagonal part. The field at any point depends on the inverse of that matrix: $\Psi(\mathbf{x})=\Psi_0(\mathbf{x})+$(vector of $f\,e^{ikr}/r$ from the scatterers to $\mathbf{x}$) $(\mathbf{I}-\mathbf{R})^{-1}$ (vector of e^{ikr}/r from the source to the scatterers). This is equivalent to Foldy's equations 7 and 8.

The multiple scattering series is obtained by the Taylor expansion in f of the inverse of the matrix,

$$(\mathbf{I}-\mathbf{R})^{-1}=\mathbf{I}+\mathbf{R}+\mathbf{R}^2+\mathbf{R}^3+\mathbf{R}^4+\cdots. \tag{2}$$

The resulting multiple scattering series has the following form of the Green's function: Each term corresponds to a sequence of scatterings; the $\mathbf{R}^n$ matrix term of the inverse has n scatterings, since every matrix element of $\mathbf{R}$ is linear in f. The same scatterer may occur more than once if there is at least one different scatterer in the sequence between any two occurrences of the same scatterer. (The absence of the same scatterer acting twice in a row follows from the absence of diagonal terms in $\mathbf{R}$.) There are propagation factors of the form e^{ikr}/r, where r is the first the distance from the source to the first scatterer, then the distance from the first to the second, and so on, and finally the distance from the last scatterer to the receiver. At each scattering event between two of these propagation factors, there is a factor of the scattering amplitude f. The Green's function is the sum of the terms for all possible sequences. (In many cases the multiple scattering series is divergent, and the result must be defined by analytic continuation in f; there are poles in f thought of as a complex variable.)

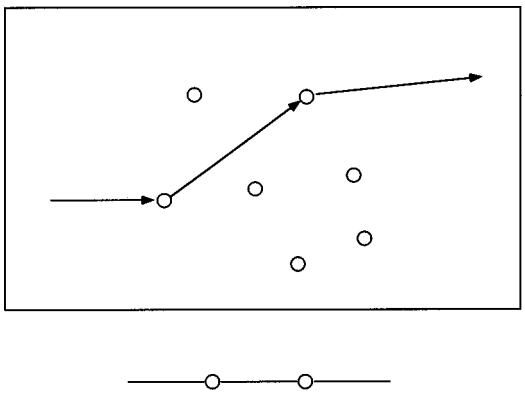


FIG. 1. Possible geometric arrangement of some scatterers with a term in the multiple scattering series indicated. Below is the diagram corresponding to that term.

Finally, the ensemble average over the positions of the scatterers must be carried out to obtain the average Green's function, e^{iKr}/r. There is an integral over the positions of the scatterers, $\Pi_{j=1}^{n}\int d^3\mathbf{x}_j\,P(\mathbf{x}_j)$, where $P(\mathbf{x}_j)$ is the probability density. For each term in the multiple scattering series, those scatterers that are not involved in that term integrate to one. There are left, for each term, as many spatial integrals as there are *different* scatterers in the sequence. The first scatterer can be any of the n scatterers that are present, the second can be any of $n-1$, etc. In keeping with Foldy's model that the scatterers are arranged as a Poisson process, we shall assume that there is also a Poisson distribution of the number of scatterers. The Poisson distribution has the property that the average combinatorial factor we need, $\langle n(n-1)\cdots(n-m)\rangle=\langle n\rangle^{m+1}$, so there is a factor of $\langle n\rangle$ for each distinct scatterer. (Alternatively, it can be imagined that n is so large that $n-1$, etc. can be replaced by n.) The probability distribution $P(\mathbf{x})$ of each scatterer is $1/V$, where V is the region occupied by the n scatterers. Thus, each integral over space effectively has a factor of $N=\langle n\rangle/V$, where N is the density of scatterers. The power of N is less than or equal to the power of f in any term.

The multiple scattering series can be presented and manipulated in a diagrammatic form. Figure 1 shows a possible geometric arrangement of scatterers with a multiple scattering term indicated. This term is then indicated by the diagram shown below the geometric picture. The diagram is both a formalized picture of the scattering term and the algebraic expression for that scattering. (Diagrams similar to these are commonly used in quantum field theory and statistical physics. Tatarskii[4] reviews the less frequent use of diagram methods in classical physics.) The diagrammatic elements are shown in Fig. 2. A line indicates the propagation factor e^{ikr}/r, for some appropriate r. One can think of the line as a ray connecting two points a distance r apart. A small circle represents a scattering amplitude f. One can think of the circle as a scatterer at some position. Circles displaced from each other represent separate scatterers; there is a factor of N and a spatial integration associated with each, due to the ensemble average over the distribution of positions. Two or more circles drawn tangent to each other indicate scatterings of the same scatterer; there is only one N and

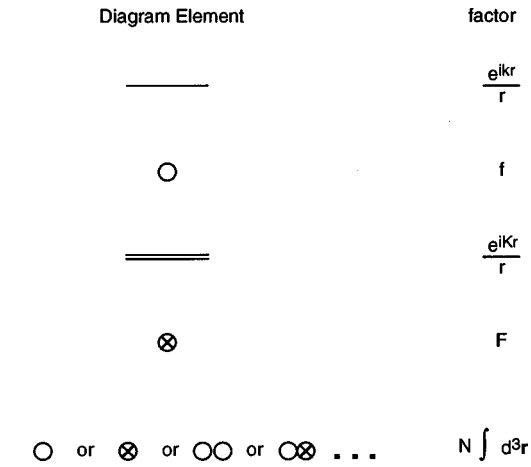


FIG. 2. Diagrammatic elements used in this paper for representing the multiple scattering series. Lines are propagation, and circles are scattering. Two or more circles drawn tangent represent more than one scattering from the same scatterer; only one spatial integral is to be done.

one integral for the group. Further diagrammatic elements, namely a double line and a small circle with an *x* in it (⊗), represent sums of contributions of more than one (possibly an infinite number) simpler diagrams. The double line represents the effective propagation e^{iKr}/r, or a model of that propagation. The circle with an x in it represents a quantity called the effective scattering amplitude, denoted by *F*, which is defined below. The reason for these additional elements is to obtain algebraic equations to be solved for *K* and *F*. Effective quantities like *K* and *F* are common in quantum field theory and statistical physics, where they are termed ''renormalized'' quantities; the approach presented here is adapted from approaches used in those other fields.

Foldy's expression is an approximate solution to his model which is given by all diagrams that have no more than one scatter from each scatterer. These diagrams are shown in Fig. 3(a). As an example of the rules for diagrams, the diagram on the second line corresponds to the expression

$$N^3 f^3 \int\int\int d^3\mathbf{x}_1\, d^3\mathbf{x}_2\, d^3\mathbf{x}_3 \times \frac{e^{ik|\mathbf{x}_1-\mathbf{x}_s|}}{|\mathbf{x}_1-\mathbf{x}_s|}\frac{e^{ik|\mathbf{x}_2-\mathbf{x}_1|}}{|\mathbf{x}_2-\mathbf{x}_1|}\frac{e^{ik|\mathbf{x}_3-\mathbf{x}_2|}}{|\mathbf{x}_3-\mathbf{x}_2|}\frac{e^{ik|\mathbf{x}_r-\mathbf{x}_3|}}{|\mathbf{x}_r-\mathbf{x}_3|}, \tag{3}$$

where $\mathbf{x}_s$ and $\mathbf{x}_r$ are the source and receiver positions.

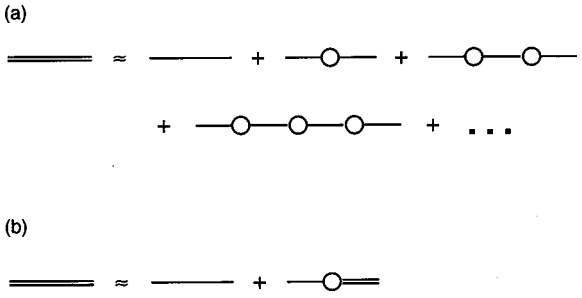


FIG. 3. (a) The sum of diagrams that give Foldy's approximate solution to his model. (b) An equation representing the same sum. All diagrams past the first in (a) consist of a line, a circle, and the previous diagram.

A convenient way to sum the diagrams of Fig. 3(a) is to convert it to an equation for *K*. One sees that each diagram, except the first, in Fig. 3(a) consists of a propagation to the first scatterer, a scattering, and the previous diagram from then on. In other words, the entire sum is the first term plus a propagation to the first scatterer, a scattering, and the sum. This is shown diagrammatically in Fig. 3(b).

The evaluation of Fig. 3(b) [or 3(a)] is most conveniently carried out in the wave number representation. They involve convolutions of e^{ikr}/r or e^{iKr}/r factors [as, for example, occurs in Eq. (3)], which become products when transformed to wave number space. The transform wave number will be denoted by *q*; the Fourier transform of e^{ikr}/r is $4\pi/(q^2-k^2)$. Thus, Fig. 3(b) reads

$$\frac{4\pi}{q^2-K^2} = \frac{4\pi}{q^2-k^2} + \frac{4\pi}{q^2-k^2} Nf \frac{4\pi}{q^2-K^2}, \tag{4}$$

which can be solved for K^2:

$$K^2 = k^2 + 4\pi Nf, \tag{5}$$

which is Foldy's result. Even this result leads to a complex wave number which represents attenuation as well as sound speed change, since *f* is complex. The argument made above that the solution must be of the form e^{iKr}/r is actually not required. One could replace $4\pi/(q^2-K^2)$ in Eq. (4) by some unknown function of *q*, find the function to be $4\pi/(q^2-K^2)$, and Fourier transform to obtain e^{iKr}/r.

It would be a mistake to assume the second diagram in Fig. 3(b) should have a double line on both sides. One might think that both the propagation to and from the scattering should be evaluated in the effective medium. However, propagation in the effective medium involves scattering. The particular scatterer exhibited could be any one of several, which would cause an overcounting. A diagram from Fig. 3(a) with *n* scatterers would be counted *n* times. The bookkeeping issue is handled in Fig. 3(b) by deciding that the very first scatterer is made explicit, and the rest are summarized in the propagation from that scatter to the receiver.

To go beyond Foldy's expression, another diagram element is used to keep track of the ''bookkeeping.'' This element is important in that the contributions it adds together are related in a way to preserve the e^{iKr}/r form of the effective propagation. In any diagram, there may be lines that if one of them is cut, that the diagram falls apart into two (each of which has at least one circle representing a scattering). A given diagram may have any number of such lines. For example, the third diagram on the right side of Fig. 3(a) can be cut apart in the middle, and the fourth can be cut apart in two places. The sum of all diagrams that cannot be so decomposed is denoted by *F*, and drawn as a circle with an x in it.

The sum of all diagrams is shown in Fig. 4(a). The resemblance to the diagrams that give Foldy's expression is obvious. The algebra proceeds in exactly the same way. Figure 4(b) shows the equation to be solved. Its solution is

$$K^2 = k^2 + 4\pi NF. \tag{6}$$

The difference is that Eq. (5) is an approximation, but Eq. (6) is exact. However, one needs an expression for *F*, which contains all the complexity of the original problem. Figure

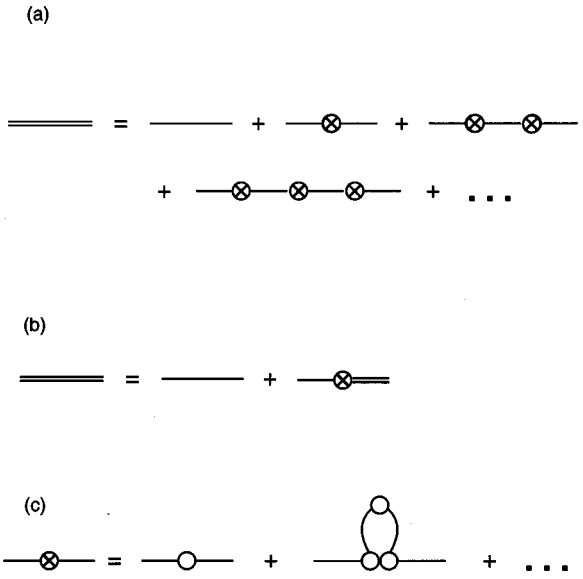


FIG. 4. (a) The sum of all diagrams. The circle with the x in it represents the sum of all diagrams that cannot be split into two pieces by cutting a single line. This is referred to as the effective scattering amplitude. (b) The equation representing the sum of the diagrams in (a). Unlike Fig. 3(b), this equation is exact. (c) The effective scattering amplitude is an infinite sum of diagrams; the first two are explicitly shown.

4(c) indicates the first two contributions to F, single scattering and a loop diagram. The loop was evaluated by Ye and Ding, giving

$$F = f + \frac{4\pi i N f^3}{2k} + \cdots . \tag{7}$$

Notice that there is no f^2 term in the expansion of F; the only f^2 term in the entire series is the third diagram on the right side of Fig. 3(a), which is not part of F because it can be cut apart. A nondiagrammatic derivation of that result is given in the Appendix.

Foldy's expression is obtained by truncating Eq. (7) at the first term, as diagrammed in Fig. 5(a). Ye and Ding's expression is obtained by truncating after the first two terms, as diagrammed in Fig. 5(b).

Foldy gave a plausibility argument for neglecting multiple scatters on the same scatterer, independent of the smallness of any parameter. However, if the two parameters Nf/k^2 and kf are both small, Foldy's expression is the lowest order for F, and Ye and Ding's is the next order. However, if

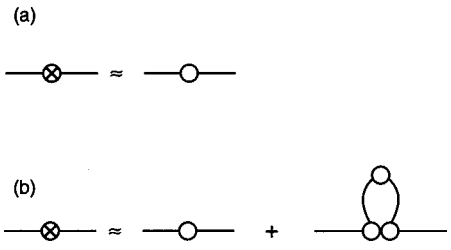


FIG. 5. (a) Foldy's approximation is that the effective scattering amplitude is the original scattering amplitude. (b) Ye and Ding's approximation consists of retaining the first two terms. When Foldy's term is large, the second term overestimates the correction, according to the argument in the text.

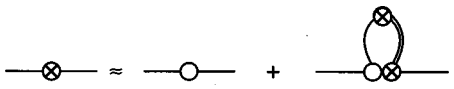


FIG. 6. The approximation advocated in the present paper. No contribution is counted more than once. By successive substitutions, diagrams with loops upon loops are included in this approximation.

these parameters are not small, the truncation after the second term can be questioned. As discussed above, it is Nf/k^2 that is not small in interesting cases.

The $1/2k$ in the second term of Eq. (7) results from the exponent in the product of the two propagations:

$$\left(\frac{e^{ikr}}{r}\right)^2 = \frac{e^{i2kr}}{r^2}. \tag{8}$$

Thus, one sees that the spatial integral is dominated by separations of the two scatters by distances on the order of $2k$. If $|K| \gg k$, the wave will not propagate that far without a high probability of undergoing further scattering. The actual dominant distance of the second scatterer should be related to K.

The model presented here takes Ye and Ding's result and combines it with other, related, diagrams, which restrict the dominant separation between the scatterers in the loop. In order that no diagram be counted more than once, the device of exhibiting only the first of a number of events is used. The model is shown in Fig. 6. The second diagram on the right side is understood as follows: The first scattering event is exhibited, so the circle has no x in it. The very next scatterer in the loop is shown, so the line to it is single. Loops of additional scattering can be attached to that scatterer, so it has an x. The further propagation back to the first scatterer can have additional scattering, so a double line is used. Finally, there can be any number of additional loops on the first scatterer after the first loop which is shown, so there is an x in the circle representing these additional scatterings. Of course, it is only the double line in that diagram which directly addresses the criticism given above of Ye and Ding's result. It turns out, however, that the physical interpretation is considerably simplified by inclusion of all these contributions.

The propagations are

$$\frac{e^{ikr}}{r}\,\frac{e^{iKr}}{r} = \frac{e^{i(k+K)r}}{r^2}, \tag{9}$$

so the $2k$ in Eq. (7) is replaced by $k+K$. Also, two of the scattering amplitudes are the effective amplitude F, and only one is f. Thus, this model gives

$$F = f + \frac{4\pi i N f F^2}{k+K}. \tag{10}$$

A factor of $4\pi NF$ from the right is replaced by $K^2 - k^2$, using Eq. (6), and the result divided by Ff, to obtain

$$\frac{1}{F} = \frac{1}{f} - i(K-k). \tag{11}$$

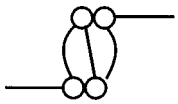


FIG. 7. The simplest diagram not included in the model of Fig. 6. This diagram is infinite, but the infinity is canceled by other diagrams; a related infinite set of diagrams gives a convergent result.

This expression can be used in Eq. (6), obtaining an equation (a cubic equation) for K. The pair of equations, (6) and (11), are the principal results of this paper.

Equation (11) has a very simple physical interpretation. As discussed by Foldy (and textbooks on wave scattering[5]), $1/f$ has an imaginary part which is the sum of contributions from various loss terms. Among the losses is the scattering loss. The scattering loss is given by an imaginary contribution of $-k$. Equation (11) says that the $-k$ is to be replaced by $-K$, the effective wave number, with no other change. The scattering loss is, therefore, to be evaluated in the effective medium, in order to obtain the effective scattering amplitude. This result was anticipated in a qualitative way by Lax,[6] who made an argument based on the optical theorem[5] to assert ''scattering amplitudes and cross sections must be evaluated for the wavelength appropriate to the medium.'' He did not give any expression for the effective scattering amplitude.

For sound propagating through water which contains small bubbles (about ten μm to several hundred μm radii), the dominant loss is thermal hysteresis. Thus a change from scattering loss in pure water to scattering loss in bubbly water is unimportant until K gets so large relative to k in order to dominate the losses.

Equation (11) is derived in a uniform medium. Clearly, one cannot use Ye and Ding's evaluation of the integral when F and K are functions of position. Since the integral is dominated by r on the order of $1/(k+K)$, presumably one should use averages over such a sphere in the correction. If F and K vary little on the scale of $1/(k+K)$, then Eq. (11), evaluated locally, should be correct.

II. OTHER CONTRIBUTIONS

We now turn to contributions not included. The simplest such diagram is shown in Fig. 7. It is infinite because the propagations give

$$\left(\frac{e^{ikr}}{r}\right)^3 = \frac{e^{i3kr}}{r^3} \tag{12}$$

and the integration over the separation, $\int (1/r^3) d^3\mathbf{r}$ is logarithmically divergent at $\mathbf{r}\to 0$.

We do not really know what probability distribution to use at short distances. If it matters, the result is uncertain. Fortunately, however, there is nothing wrong with Foldy's model in connection to the diagram of Fig. 7. In order to see that there is no problem, consider diagrams like Fig. 7, but with 5,7,9,... lines connecting the two scatterers. These have factors $1/r^5, 1/r^7, 1/r^9, \ldots$, all of which give divergent integrals at $\mathbf{r}\to 0$, getting more divergent as the order of the diagram increases. The quantity of interest is the mean field; it is not necessary that the mean of a part of the field exist. Thus, the summation of diagrams is supposed to be done before the integration over the probability distribution of the positions of the scatterers; only when it does not matter (and it has not until Fig. 7 was considered) can we interchange the order of the two operations. Therefore we should first sum the set of related diagrams. Each higher diagram of the set has two extra scattering factors of f and two extra propagation factors e^{ikr}/r. The sum gives the first diagram of the set multiplied by

$$\sum_{j=0}^{\infty} f^{2j} \frac{e^{2ijkr}}{r^{2j}} = \frac{1}{1-(f(e^{ikr}/r))^2}.$$

This factor vanishes quadratically at $r=0$, making the integral convergent. The cutoff provided by this factor is at $r \approx |f|$.

It is likely that every divergent diagram is part of a set, and the sum of the set is finite after integration for a reason similar to the set associated with Fig. 7. Each cluster of very nearby scatterers acts as a resonator, and the pole associated with the resonance is at a complex frequency, although it is manifested as singularities at spacing$=0$ in individual diagrams. If that is the case, no modification of Foldy's model need be made, and his approximation that scatterers can get arbitrarily close together causes no harm, and if the scatterers are small, can have very little loss of accuracy.

III. RELATED WORK

Other recent work on this topic is that of Sangani[7] and that of Feuillade.[8] Sangani does not work in the framework of Foldy's model. In fact, he attributes to Foldy an expression which is far simpler than Eq. (1). In particular, Sangani's version of Foldy's result has a real expression in place of f; Foldy devotes an entire section of his paper explaining the importance of the imaginary part of f. As a result, much of what Sangani presents as new is already part of Foldy's result. Sangani modifies Foldy's model by introducing a cutoff at the size of the scatterer. Sangani finds an important contribution from a term which is represented by Fig. 7; the term is dominated by the cutoff.

The cutoff provided by higher-order terms is at $r \approx |f|$, which can be considerably larger than the size of a scatterer, especially when the scatterer resonates. In such a case, Sangani's treatment of this term overestimates the effect.

Sangani proposed an expression (the final form in his equation 5.23) which is equivalent to Eq. (11) above, and discusses the physical interpretation. No derivation is provided other than the calculation of the first two terms in what amounts to the Taylor series in K of F. This result is supposed to hold for $K \gg k$, for which his first form agrees with Ye and Ding, which he then equates to the final form.

Feuillade has proposed a correction to Foldy's result based on intuitive arguments. His correction will be shown to be incorrect. Most significantly, $\mathbf{p}$ is defined below his equation 11 as the set of values of the external pressure at each bubble (the aggregate of the scattered fields is added to get the total pressure). Following his equation 29 he states that he interprets its average as the ''average pressure field experienced by the bubbles... .'' He intends (private commu-

nication) this to mean the average external pressure. However, the compressibility is expressed in terms of the total average pressure field experienced by the bubbles, which also includes that of the sound scattered by other bubbles. If one corrects his result for that effect, his result turns out to be identical to Foldy's. This is not surprising, as Foldy's expression accounts for those collective effects that result from averaging the relative positions of the scatterers in each interaction separately. Less significantly, in his equation 11 (which is asserted without derivation), the wave number k must be the wave number in the original medium, as derived by Foldy (Foldy's equation 8). Feuillade (private communication) intends it to be the wave number in the effective medium, as is consistent with the remainder of his paper. Even without these specific identifications of his errors, one can see that his result is incorrect, since a Taylor series expansion of his result gives an order f^2 correction to Eq. (1), whereas Ye and Ding have rigorously shown the leading correction to be of order f^3. The absence of the f^2 correction is demonstrated in detail in the Appendix.

IV. SUMMARY

A correction to Foldy's effective medium expression has been derived. This correction is that the scattering loss should be evaluated in the effective medium; otherwise, the expression is the same as Foldy's. Additional contributions become important when the average spacing of scatterers is on the order of the absolute value of the scattering amplitude. The correction of order $N|f|^3$ relative to Foldy's term has not been estimated, so the results require that parameter to be small. Evaluation of the $N|f|^3$ terms require that the multiple scattering series be summed before the integration over scatterer positions, in order to avoid a divergent answer.

Previous results are expressible in the form of Eq. (6). Foldy has $F=f$, which is the first term in the power series for F; Ye and Ding have $F=f+4\pi iNf^3/2k$, which are the first two terms; Feuillade has $F=f/(1-4\pi NfI)$ (see the Appendix), which contains a spurious quadratic term; the present work has $F=f+4\pi iNfF^2/(k+K)$, which agrees with Ye and Ding through $O(f^3)$, but also includes contributions that nearly cancel the cubic term when K is large compared to k, giving a result very close to Foldy's.

ACKNOWLEDGMENTS

I would like to thank my colleagues at the Applied Physics Lab, in particular R. Odom, S. Kargl, D. Jackson, and E. Thorsos for reading a draft of the manuscript and suggesting improvements. This work is supported by the Office of Naval Research, Code 321OA.

APPENDIX: ABSENCE OF A QUADRATIC CORRECTION TO FOLDY'S RESULT

In this Appendix, we compare Feuillade's expression through quadratic order with the actual multiple scattering series, and show they are in disagreement. The algebra is carried out without the use of diagrams.

Feuillade's equation 30, when expressed in terms of the variables (and phase convention) used here, is

$$K^2=k^2+\frac{4\pi Nf}{1-4\pi NfI}, \tag{A1}$$

where $I=\int_0^\infty re^{ikr}\,dr$. The Taylor series of Eq. (A1) is

$$K^2=k^2+4\pi Nf+(4\pi Nf)^2I+O(f^3), \tag{A2}$$

which has a nonzero quadratic part.

The actual multiple scattering is given by

$$\frac{e^{iKr}}{r}=\frac{e^{ikr}}{r}+Nf\int d^3\mathbf{x}_1\frac{e^{ik|\mathbf{x}_1-\mathbf{x}_s|}}{|\mathbf{x}_1-\mathbf{x}_s|}\frac{e^{ik|\mathbf{x}_r-\mathbf{x}_1|}}{|\mathbf{x}_r-\mathbf{x}_1|}+N^2f^2 \times\int\int d^3\mathbf{x}_1\,d^3\mathbf{x}_2\frac{e^{ik|\mathbf{x}_1-\mathbf{x}_s|}}{|\mathbf{x}_1-\mathbf{x}_s|}\frac{e^{ik|\mathbf{x}_2-\mathbf{x}_1|}}{|\mathbf{x}_2-\mathbf{x}_1|}\frac{e^{ik|\mathbf{x}_r-\mathbf{x}_2|}}{|\mathbf{x}_r-\mathbf{x}_2|} +O(f^3). \tag{A3}$$

The integrals evaluate to

$$\frac{e^{iKr}}{r}=\frac{e^{ikr}}{r}\left(1+\frac{2\pi iNfr}{k}+\frac{2(\pi iNf)^2r}{k^2}\left(r+\frac{i}{k}\right) +O(f^3)\right). \tag{A4}$$

Multiplying by r and taking the logarithm leads to

$$iKr=ikr+\frac{2\pi iNfr}{k}-\frac{2i(\pi Nf)^2r}{k^3}+O(f^3). \tag{A5}$$

Dividing by ir and squaring gives

$$K^2=k^2+4\pi Nf+O(f^3). \tag{A6}$$

There is no quadratic term, which contradicts Feuillade's result. The cancellation appears coincidental in the presentation of this Appendix, but occurs very naturally in the diagrammatic method.

[1] L. L. Foldy, ''The multiple scattering of waves. I. General theory of isotropic scattering by randomly distributed scatterers,'' Phys. Rev. **67**, 107–119 (1945).

[2] Z. Ye and L. Ding, ''Acoustic dispersion and attenuation relations in bubbly mixture,'' J. Acoust. Soc. Am. **98**, 1629–1636 (1995).

[3] E. Terrill and W. K. Melville, ''Sound-speed measurements in the surface-wave layer,'' J. Acoust. Soc. Am. **102**, 2607–2625 (1997).

[4] V. I. Tatarskii, ''The effects of the turbulent atmosphere on wave propagation,'' Israel Program for Scientific Translation (1971), Chap. 5.

[5] Historical and textbook references are given in R. G. Newton, ''Optical theorem and beyond,'' Am. J. Phys. **44**, 639–642 (1976).

[6] M. Lax, ''Multiple Scattering of Waves,'' Rev. Mod. Phys. **23**, 287–310 (1951).

[7] A. S. Sangani, ''A pairwise interaction theory for determining the linear acoustic properties of dilute bubbly liquids,'' J. Fluid Mech. **232**, 221–284 (1991).

[8] C. Feuillade, ''The attenuation and dispersion of sound in water containing multiply interacting air bubbles,'' J. Acoust. Soc. Am. **99**, 3412–3430 (1996).

On bistatic sea surface scattering: Field measurements and modeling

Peter H. Dahl
Applied Physics Laboratory, College of Ocean and Fishery Sciences, 1013 N.E. 40th Street, Seattle, Washington 98105-6698

(Received 9 March 1998; revised 6 November 1998; accepted 11 December 1998)

The bistatic scattering cross section of the sea surface, σ, is studied, along with a model for σ and its comparison with field data. The data are horizontal spatial coherence and ensemble-averaged intensity, which represent integral measures of sea surface bistatic scattering, and the model for σ is used to generate these same properties for comparison with the field data. The data are from an experiment conducted in shallow waters off southern Florida, using a sound frequency of 30 kHz. Directional wave measurements were made with a wave buoy positioned within 100 m of the acoustic measurements, with the environment characterized by rms wave heights of $O(10)$ cm and wind speeds of 1–4 m/s. In the model σ is divided into two components: σ_r associated with scattering from the rough, air/sea interface, and σ_b associated with scattering from near-surface bubbles. The second-order small slope approximation is used to compute σ_r, which is a much improved approach over the traditionally used composite roughness model. The primary advantage in the small slope approximation was the resulting smooth behavior in σ_r over a broad range of scattering angles. Directional wave data obtained by the wave buoy were converted to an estimate of the 2-D spatial correlation function of sea surface roughness, $C(\xi,\varsigma)$, for use in the scattering calculations. An analysis of the effective correlation properties of $C(\xi,\varsigma)$ suggested that an isotropic correlation function $C(\rho)$, based on the directionally averaged wave-number spectrum, would be equally effective in the scattering calculations. Model-data agreement was quite satisfactory, regardless of whether $C(\xi,\varsigma)$ or $C(\rho)$ was used in the scattering calculations. © *1999 Acoustical Society of America.* [S0001-4966(99)01904-9]

PACS numbers: 43.30.Hw, 43.30.Re, 43.30.Zk, 43.30.Gv [DLB]

INTRODUCTION

Bistatic scattering from the sea surface is a key descriptor of the acoustic environment, and strongly influences the performance of sonar systems that operate either in shallow water or in the vicinity of the sea surface. In this paper, measurements of sound that has undergone bistatic forward scattering from the sea surface are presented, together with a new model for the sea surface bistatic cross section, σ. The model includes scattering from both the rough air/sea interface and near-surface bubbles. For roughness scattering we use the small slope approximation originally developed by Voronovich.[1] Our primary motivation for using the small slope approximation is that it eliminates the need to match cross sections derived from the perturbation and Kirchhoff approximations.

The data are in the form of (a) horizontal spatial coherence across a line array oriented transverse to the direction of propagation and (b) calibrated level and temporal decay of ensemble-averaged intensity for a single receiving element. These represent *integral* measures of the sea surface bistatic scattering. The model for σ is used to generate these same properties for comparison with the field data. For the model calculations, we incorporate directional wave measurements that were taken simultaneously with the acoustic measurements, and we also examine the effects of surface wave directivity on bistatic forward scattering.

The measurements are from an experiment conducted in shallow waters off South Florida. The sound frequency was 30 kHz, and the air/sea conditions were characterized by wind speeds of 1 to 4 m/s and rms waveheights of 8 to 16 cm (significant waveheights of 32 to 64 cm). Note that this study concerns only incoherent scattered intensity (on an ensemble mean basis) from the rough air/sea interface and underlying bubble layer, the coherent scattered intensity being negligible for this frequency range in all but mirrorlike sea surface conditions. However, spatial coherence, which depends on the relative phase difference between sensors on an array, still has meaning. Horizontal spatial coherence relates directly to the horizontal spread in the direction of arrivals at the array and sets the effective aperture length of the array for directional discrimination.

The relationship between horizontal spatial coherence and the bistatic scattering cross section was originally presented by Dahl[2] (see Sec. II of that reference) who also discusses results of a similar experiment but performed in a pelagic environment. The link between the two is based on forming a probability density function for bistatic scattered intensity versus horizontal arrival angle, which is Fourier transformed to give a model for the horizontal spatial coherence. Thus, the approach outlined in Ref. 2 highlights the interrelations between the bistatic cross section, the horizontal angular density function, and the horizontal spatial coherence. Note that this present work replaces that discussed in Ref. 2 insofar as the model used for the bistatic cross section.

The experiment and analysis methods are described in Sec. I, and the components of the model for the bistatic cross

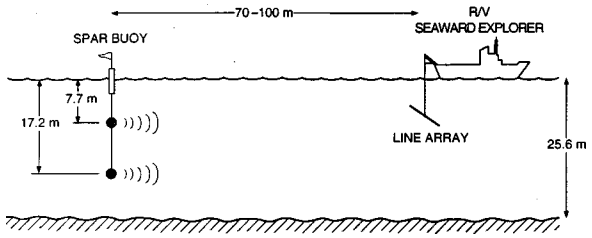


FIG. 1. Experimental layout for measuring bistatic forward scattering from the sea surface.

section are discussed in Sec. II. In Sec. III a brief study on sea surface bistatic scattering as it relates to the directional wave spectrum and near-surface bubbles is presented along with model computations of the bistatic cross section σ. The model results for spatial coherence and intensity and their comparison with field measurements are given in Sec. IV, and a summary is given in Sec. V. In Appendix A we discuss the surface waveheight spectral measurements made with a directional wave buoy, which are incorporated into the scattering calculations, and Appendix B contains a derivation for a correction term that accounts for the tilt modulation of large-scale waves.

I. EXPERIMENTAL DESCRIPTION

The experiment was conducted 13–16 February 1995 near the Dry Tortugas collection of islands (24°36.7′N 82°50.7′W) off southern Florida. The water depth was 25.6 m, and the bottom sediment consisted of carbonate sand–silt–clay.[3] The experimental layout for the bistatic forward-scattering measurements is shown in Fig. 1. Signals were transmitted from one of two ITC-1032 omnidirectional transducers suspended from a spar buoy at depths of 7.7 and 17.2 m. The spar buoy was tethered to the research vessel SEAWARD EXPLORER (which held position in a four-point moor) using a distributed-buoyancy power line, at ranges between 70 and 100 m; a small boat was used for final positioning. The natural period of the spar buoy was 12 s, specifically designed to be well away from the period of the dominant surface waves which was about 6 s.

An experimental run consisted of either 10 or 20 pulses transmitted at 5-s intervals. The pulses were cw transmissions of length 1 ms and frequency 30 kHz. The acoustic measurements are organized into four runs (Table I). Run 1 was taken 3 days before runs 2–4, with runs 2 and 3 taken sequentially and run 4 taken 50 min later. The signals were received on a horizontal line array suspended off the stern of the research vessel at a depth between 7 and 10 m. A bracing structure was used to restrict uncontrolled array rotation, and a damping mechanism was used to reduce array heave motion. We simultaneously recorded signals from the eight-element array, on which the minimum and maximum element separations were 1.59 and 57.15 cm, respectively. (Note that during the experiment, elements 7 and 8 failed, which reduced the maximum separation to 31.75 cm and the element-spacing combinations from 28 to 15.) Data-reduction methods and the procedure for estimating spatial coherence are the same as those described in Ref. 2. This procedure also includes a phase correction for any slight rotation of the array, which invariably occurs during the 100-s averaging interval. We assume the received signal (of relative bandwidth 3%) is composed of only one frequency component; thus the results for spatial coherence are expressed in terms of the receiver separation normalized by the acoustic wavelength equivalent to 30 kHz.

Wind speed measurements were made using a Coastal Climate propeller vane anemometer (from which air temperature and humidity were also recorded) mounted 6.9 m above the sea surface at the top of the hydraulic U-frame located on the stern end of the *Seaward Explorer*. This location resulted in minimal interference from vessel structures. Wind speed was recorded continuously as a 5-min average every 10 min. Surface-wave directional spectra were measured using a Datawell Directional Waverider buoy (from which sea temperature was also recorded). The buoy was moored approximately 100 m from the *Seaward Explorer*, and data were telemetered back to the vessel every hour. The wave buoy data are discussed in more detail in Appendix A. Finally, conductivity, temperature, and pressure (CTD) measurements were made approximately every 3 h using a Sea-Bird CTD unit. These data showed the water column to be well mixed down to bottom, with an average sound speed of 1520 m/s.

Figures 2(a) and 3(a) illustrate the dominant (i.e., specular) ray paths that compose the arrival structure for experimental runs 1 and 2, respectively. With a 1-ms pulse, the first few paths can be resolved and their properties studied separately. For example, Figs. 2(b) and 3(b) show the time series of ensemble-averaged intensity, for which the surface-bounce, bottom-bounce, and surface–bottom bounce paths are easily resolved. Our focus in this paper is solely on the single surface-bounce paths, and results of an initial study on bottom-plus-surface interacting paths are described by Dahl and Fox.[4] Notice, however, that paths involving the bottom suffer an energy loss of 20 to 25 dB because the bottom grazing angle exceeded the critical angle (approximately 10 degrees). On the other hand, the single surface-bounce path in each case experiences essentially 0 dB energy loss. Thus,

TABLE I. Key variables defining scattering geometry and conditions (columns 2–9) and estimated parameters (columns 10 and 11). See text for definitions.

Run	Source depth (m)	Receiver depth (m)	Range (m)	θ_g (deg)	Wind speed (m/s)	Wave height, rms (cm)	χ	s_L	$kd^\star/(2\pi)$	σ_{θ_h} (deg)
1	17.2	6.9	107	13	$\lesssim$1	8	4.4	0.060	20.0	0.45
2	7.7	9.5	72	14	4	16	9.5	0.130	12.0	0.76
3	17.2	9.5	72	20	4	16	13.4	0.130	4.9	1.87
4	7.7	10.0	75	13	3.8	16	9.1	0.134	11.2	0.82

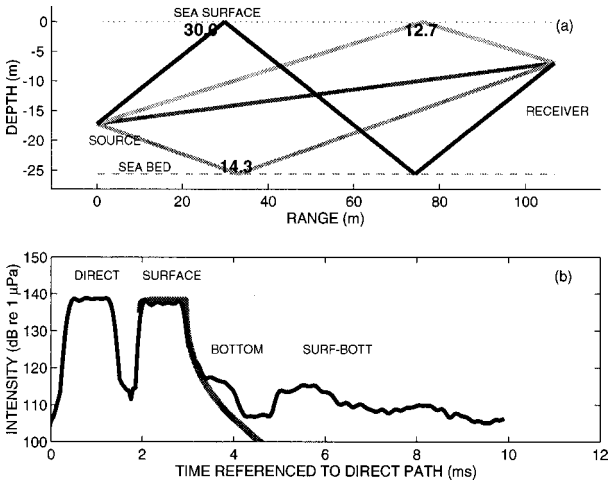


FIG. 2. (a) Dominant (specular) ray paths corresponding to experimental run 1. (b) Ensemble-averaged intensity of signal received on an array element showing the primary arrivals: direct, surface bounce, bottom bounce, and surface–bottom bounce. Solid curve plotted over the surface bounce segment of the data is a model result discussed further in Sec. IV.

for this combination of $O(100)$ m range and mid-water source and receiver depths, the acoustic channel is dominated by the direct and single-surface-bounce paths. We postpone discussion of the model curves plotted over the ensemble-averaged intensity data [one in Fig. 2(b), and two in Fig. 3(b)] until Sec. IV.

II. COMPONENT MODELS FOR THE BISTATIC SCATTERING CROSS SECTION OF THE SEA SURFACE

In this section we present expressions for the bistatic cross section of the sea surface, σ, which we then use to generate integral properties of bistatic surface scattering for comparison with field data. We divide σ into two components: σ_r, which is due to scattering from the rough air/sea interface, and σ_b, which is due to scattering from near-surface bubbles, with σ equal to $\sigma_r+\sigma_b$ (incoherent addition). The discussion below will initially focus on the approach to approximating σ_r.

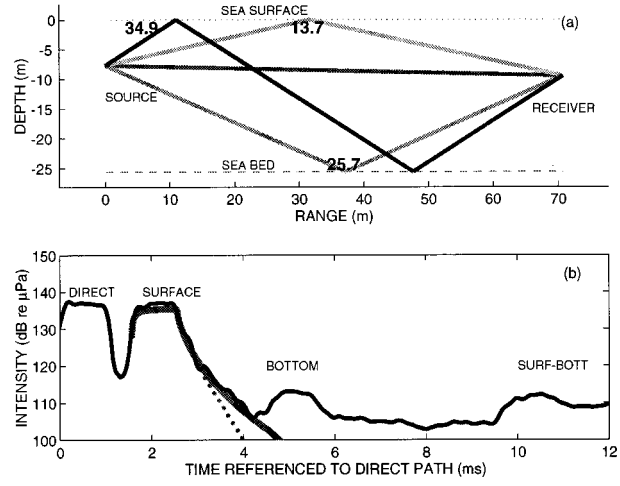


FIG. 3. (a) Dominant (specular) ray paths corresponding to experimental run 2. (b) Ensemble-averaged intensity of signal received on an array element showing the primary arrivals: direct, surface bounce, bottom bounce, and surface–bottom bounce. Solid and dotted curves plotted over the surface bounce segment of the data are model results discussed further in Sec. IV.

The scattering contribution from the rough air/sea interface has traditionally been estimated by dividing the surface waveheight spectrum into two parts: a large-scale component, consisting primarily of gravity waves whose wavelength is long compared with the acoustic wavelength, and a small-scale component equal to the remaining, high wave number, portion of the spectrum which consists of both gravity-capillary and capillary waves. Scattering from within the near-specular angular range is handled by the Kirchhoff approximation involving large-scale waves, whereas scattering away from the specular range is handled by the first-order perturbation approximation, which involves small-scale, or Bragg, waves. Furthermore, cross sections derived from first-order perturbation theory are modified by weighting them over a probability density function of large-scale slopes. The combination of these two approximations in the above manner is known as the composite-roughness model used in both microwave and acoustic scattering models.[5–8] This approach has been extremely useful for interpreting measurements of both microwave and acoustic scattering from the sea surface. Its primary drawback is that cross sections derived from the perturbation and Kirchhoff approximations must be matched over some angular transition range that also depends on scattering angle; this is particularly important for the case of bistatic scattering. McDaniel[10] describes a potentially useful method that achieves a smooth transition between the perturbation and Kirchhoff approximations, using correction terms obtained by applying the composite roughness model to an expansion of the scattering integral. However, the method still requires division of the surface waveheight spectrum, which is a procedure we ultimately seek to avoid in handling the transition problem.

Thus, in this paper we use the small slope approximation (SSA) to compute the bistatic cross section of the sea surface. Using this approach, a smooth transition between the perturbation and Kirchhoff approximations is achieved without requiring division of the surface waveheight spectrum into large-scale and small-scale parts. Dashen *et al.*[11] analyze the validity of the SSA and how it relates to other modeling approaches in the context of acoustic scattering from the ocean surface. Thorsos and Broschat[12] provide a derivation of the SSA subject to the Dirchlet boundary conditions, and Broschat and Thorsos[13] study the behavior of the SSA *vis-à-vis* the perturbation and Kirchhoff approximations, and exact numerical methods for scattering from 1-D surfaces characterized a Gaussian spectra. Their study demonstrates how the lowest-order SSA is accurate over a broad range of scattering angles, reducing to either the first-order perturbation approximation or the Kirchhoff approximation within the range of scattering angles where each is considered accurate. In a more recent work, Thorsos and Broschat[14] illustrate the acceptable accuracy of the lowest-order SSA (again through comparison with exact calculations) for bistatic scattering from 1-D surfaces characterized by a Pierson–Moskowitz spectrum. This spectrum exemplifies a power law spectrum

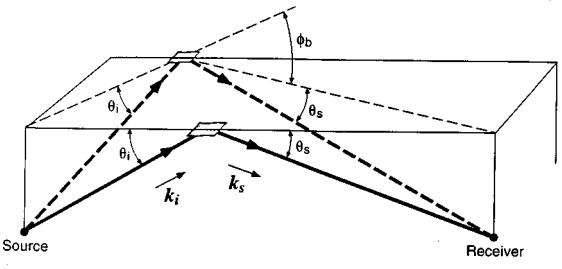


FIG. 4. Sketch showing the scattering geometry and key variables. Two scattering patches on the sea surface are shown, one representing in-plane scattering (for which the bistatic angle ϕ_b is zero) and one representing out-of-plane scattering.

similar to the measured sea surface spectra we use in this study.

It is convenient to start with the familiar Kirchhoff approximation for the incoherent bistatic cross section, σ_{KA}, given by (e.g., see Ref. 15)

$$\sigma_{\mathrm{KA}}=\frac{k^2\mathcal{F}^2}{(2\pi)^2}\int_{-\infty}^{\infty}d\xi\int_{-\infty}^{\infty}d\varsigma\, e^{2ik(\alpha\xi+\beta\varsigma)}e^{-\chi^2[1-C(\xi,\varsigma)]}, \tag{1}$$

where $C(\xi,\varsigma)$ is the autocorrelation function of surface waveheight variation and χ is the surface roughness, or Rayleigh, parameter defined as

$$\chi=kH(\sin\theta_i+\sin\theta_s), \tag{2}$$

and where k is acoustic wave number, H is rms surface roughness, and θ_i and θ_s are the incident and scattered grazing angles, respectively (Fig. 4). The parameters (α,β,γ) are the components of the vector $(\mathbf{k}_i-\mathbf{k}_s)/k$, where $\mathbf{k}_i$ is the incident wave vector and $\mathbf{k}_s$ is the scattered wave vector, and are given by

$$\begin{aligned}\alpha&=(\cos\theta_i-\cos\theta_s\cos\phi_b)/2,\\ \beta&=-(\cos\theta_s\sin\phi_b)/2,\\ \gamma&=-(\sin\theta_i+\sin\theta_s)/2,\end{aligned} \tag{3}$$

where ϕ_b is the bistatic angle and $\mathcal{F}$ equals $(\alpha^2+\beta^2+\gamma^2)/\gamma$. In our calculations we align the axis of the lag variable ξ with the acoustic line-of-sight axis along the surface, which makes the bistatic cross section a function of three angles, θ_i, θ_s, and ϕ_b, rather than four (i.e., a second bistatic angle). Note that we are not assuming isotropy at this stage as, in general, we must generate a new $C(\xi,\varsigma)$ to evaluate a different source-to-receiver direction.

Let us also define $C(\rho)$ as the equivalent isotropic autocorrelation function of surface waveheight variation which is a function of only the radial lag coordinate ρ. Then, in this case σ_{KA} is given by[15,16]

$$\sigma_{\mathrm{KA}}=\frac{k^2\mathcal{F}^2}{2\pi}\int_0^{\infty}\rho J_0(2\kappa\rho)e^{-\chi^2[1-C(\rho)]}\,d\rho, \tag{4}$$

where J_0 is the zeroth-order Bessel function, and κ equals $k\sqrt{\alpha^2+\beta^2}$. We shall refer to Eqs. (1) and (4) as the 2-D and 2-D isotropic Kirchhoff integrals, respectively, and compute both versions in order to study the effects of surface-wave directivity on our bistatic forward-scattering measurements. Looking ahead, we shall find that the difference between results derived from Eqs. (1) and (4) are, in fact, quite small when applied to our measured wave conditions and acoustic frequency of 30 kHz.

Next we convert σ_{KA} to the second-order SSA cross section, or $\sigma_{\mathrm{SSA(2)}}$, and it is this cross section that we use for our estimate of σ_r. (The order is expressed in terms of generalized slope as discussed in Ref. 12, and $\sigma_{\mathrm{SSA(2)}}$ is the lowest-order SSA.) Conversion is done by way of Eq. (15) of Ref. 13. We make a minor change in this equation so it applies to our case of 3-D scattering from a 2-D surface, with the modified equation being

$$g=\frac{\gamma\sin\theta_i-\alpha\cos\theta_i}{\sin\theta_i\sin\theta_s}, \tag{5}$$

with

$$\sigma_{\mathrm{SSA(2)}}=\sigma_{\mathrm{KA}}/g^2. \tag{6}$$

One revealing property of the SSA(2) is easily seen upon evaluating Eq. (4), together with Eqs. (5) and (6), for $\chi\ll 1$. The result, correct to order $(kH)^2$, is identical to the incoherent scattering cross section from first-order perturbation theory, which we designate as σ_{PT},

$$\sigma_{\mathrm{PT}}=4k^4[\sin^2\theta_i\sin^2\theta_s]\frac{F(K=2\kappa)}{2\pi}. \tag{7}$$

This result is in fact inherent to the construction of the small slope approximation.[13] In Eq. (7), F is the directionally averaged wave-number spectrum (Appendix A), which is a function of surface wave number K and is normalized such that

$$H^2=\int_0^{\infty}KF(K)\,dK, \tag{8}$$

and F is evaluated at the Bragg wave number equal to 2κ. For the equivalent limit of Eq. (1), $F(2\kappa)/2\pi$ is replaced by $W(2k\alpha,2k\beta)$, where W is the surface-wave-vector spectrum and the Bragg wave vector has components $(2k\alpha,2k\beta)$ matching the horizontal components of the vector $\mathbf{k}_i-\mathbf{k}_s$.

Another property of the SSA(2) is seen by studying the behavior of Eq. (5) for scattering angles approaching specular: g approaches 1 and equals 1 at the precise specular angle, which recovers σ_{KA} exactly. Thus, the aforementioned problem of matching cross sections derived from the perturbation and Kirchhoff approximations is eliminated by using the small slope approximation for all scattering angles.

In practice, however, there are numerical tradeoffs to consider in applying the small slope approximation, and elements of the composite roughness model remain very useful. For example, within the scattering angular range equivalent to the Bragg scattering regime [i.e., where Eq. (7) is considered accurate], the Kirchhoff integrand becomes progressively more oscillatory with increasing Bragg wave number. Under these conditions numerical integration can yield an erroneous result unless extreme care is taken to resolve the integrand and properly account for its cancellations (see Ref. 11 for a discussion of the numerical issues involved). Careful integration of Eq. (1) or (4) within the Bragg scattering regime, however, still does not include the

effects of tilting of the large-scale surface upon which the Bragg waves propagate (tilt modulation) nor does it include the advection of Bragg waves by the large-scale surface (hydrodynamic modulation).[17] The effect of tilt modulation appears to be sufficiently incorporated by the third-order small slope approximation, or SSA(3), as indicated by the numerical studies described in Ref. 13, although evaluation of SSA(3) is considerably less tractable than that of SSA(2).

Perturbation theory, on the other hand, lends itself to a relatively simple correction within the framework of the composite roughness model. Large-scale effects are readily dealt with by correction terms derived by carrying out the expectation of either Eq. (7) in its entirety, or of the geometric factor $\sin^2\theta_i \sin^2\theta_s$, over a probability density function for the large-scale slope. Plant,[17] for example, provides correction terms for tilt and hydrodynamic modulation for radar backscattering, for which the primary controlling parameter is the mean square slope of the large-scale waves, or s_L^2.

Thus it remains advantageous to preserve the scale separation idea in the composite roughness model, if only for the purpose of computing correction terms that account for large-scale effects on scattering in the Bragg regime. In this work, we define a separation wave number K_{sep} equal to $0.25k$, as suggested by the numerical simulations of McDaniel and McCamman,[18] and compute s_L^2 via

$$s_L^2 = \int_0^{K_{\text{sep}}} K^3 F(K)\, dK. \tag{9}$$

This separation rule is in accord with $K_{\text{sep}} \approx 0.20k$ used in radar backscattering from the sea surface.[7,17] It has been shown, for example, that radar-derived mean-square slope estimates are sensitive only to waves longer than about five times the microwave length.[19] For realistic sea surface spectra, the difference between using $0.25k$ and $0.20k$ in computing s_L^2 is nearly transparent. As a practical matter, the forward-scattering geometries involved in this experiment, plus the small values estimated for s_L^2, are such that the effect of large-scale waves on scattering in the Bragg regime becomes insignificant. However, for more general bistatic scattering geometries and conditions, a correction term to be added to σ_{PT} that accounts for tilt modulation is $4k^4 T_1 F(2\kappa)/2\pi$, where

$$\begin{aligned} T_1 = s_L^2 [& \tfrac{1}{2}\sin^2\theta_i \cos^2\theta_s + \tfrac{1}{2}\sin^2\theta_s \cos^2\theta_i \\ & - 2\sin\theta_i \cos\theta_i \sin\theta_s \cos\theta_s \cos\phi_b \\ & - \sin^2\theta_i \sin^2\theta_s]. \end{aligned} \tag{10}$$

This term is derived in Appendix B.

We leave out the correction for hydrodynamic modulation, which is significantly less than that for tilt modulation, as we expect it would not be observable in typical field data, particularly when the contribution from near-surface bubbles is added. We also proceed for now on the assumption that shadowing, which could play a role at extremely shallow incident and scattered grazing angles, can be ignored, as its principal effect will be obscured by bubble scattering.

Finally, scattering and attenuation from subsurface bubbles contributes to the total bistatic cross section σ in the following way:[2]

$$\sigma = \sigma_r \alpha_b + \sigma_b . \tag{11}$$

The factor α_b in the first term accounts for attenuation and is given by

$$\alpha_b = e^{-(\beta_I/\sin\theta_i + \beta_I/\sin\theta_s)}, \tag{12}$$

where β_I is a dimensionless parameter that succinctly describes the bubble concentration and is equal to the depth-integrated extinction cross section per unit volume. The second term is the additional isotropic scattering from bubbles and is given by

$$\begin{aligned} \sigma_b = \frac{\beta_I}{4\pi}\frac{\delta_r}{\delta}\Bigg[& \frac{1 - e^{-2(\beta_I/\sin\theta_i + \beta_I/\sin\theta_s)}}{\beta_I/\sin\theta_i + \beta_I/\sin\theta_s} \\ & + \frac{e^{-2\beta_I/\sin\theta_i} - e^{-2\beta_I/\sin\theta_s}}{\beta_I/\sin\theta_s - \beta_I/\sin\theta_i}\Bigg], \end{aligned} \tag{13}$$

where δ_r is the radiation damping coefficient at resonance, taken to be 0.0136, and δ is the total damping coefficient at resonance, equal to approximately 0.0792 at 30 kHz. Equations (12) and (13) are discussed in the context of spatial coherence in Ref. 2 [Eq. (12) is printed with an error in that reference], and the mononstatic version of these equations is discussed in the context of high-frequency surface backscattering measurements by Dahl *et al.*[20]

III. A STUDY OF THE SEA SURFACE BISTATIC CROSS SECTION RELATING TO THE DIRECTIONAL WAVE SPECTRUM AND NEAR-SURFACE BUBBLES

In this section we carry out a brief study on the sea surface bistatic scattering cross section σ as it relates to the directional wave spectrum (in terms of the total wave variance, wavelength scales, and directivity), and near-surface bubbles. We compute σ at 30 kHz frequency using the SSA(2), for two environmental conditions representing our field experiment. In one case we also compute σ using first-order perturbation theory for an important comparison.

It is instructive to first study the behavior of the integrand in Eq. (1) with respect to the correlation function $C(\xi,\varsigma)$ estimated from the wave buoy data (Appendix A). We use an estimated $C(\xi,\varsigma)$ that is a low-pass version of the true correlation function insofar as it is supported only by length scales whose wave numbers are $\leqslant K_{\text{sep}}$, since our main purpose here is to see how the integrand behaves in the near-specular scattering region, where the scattering is insensitive to small-scale surface features. The integrand is effectively controlled by the factor $e^{-\chi^2[1-C(\xi,\varsigma)]}$, which we call a *visibility function* for the Kirchhoff integral. Let us find the distance $\sqrt{\xi^2+\varsigma^2}$ at which the visibility function is reduced to 0.1, or a reduction of 10 dB, and call this an effective correlation distance for near-specular scattering. We choose χ corresponding to the specular point as representative, and the effective correlation distance for experimental run 2 (Table I) for which $\chi = 9.5$ is shown traced over the corresponding estimate of $C(\xi,\varsigma)$, forming the white dashed line

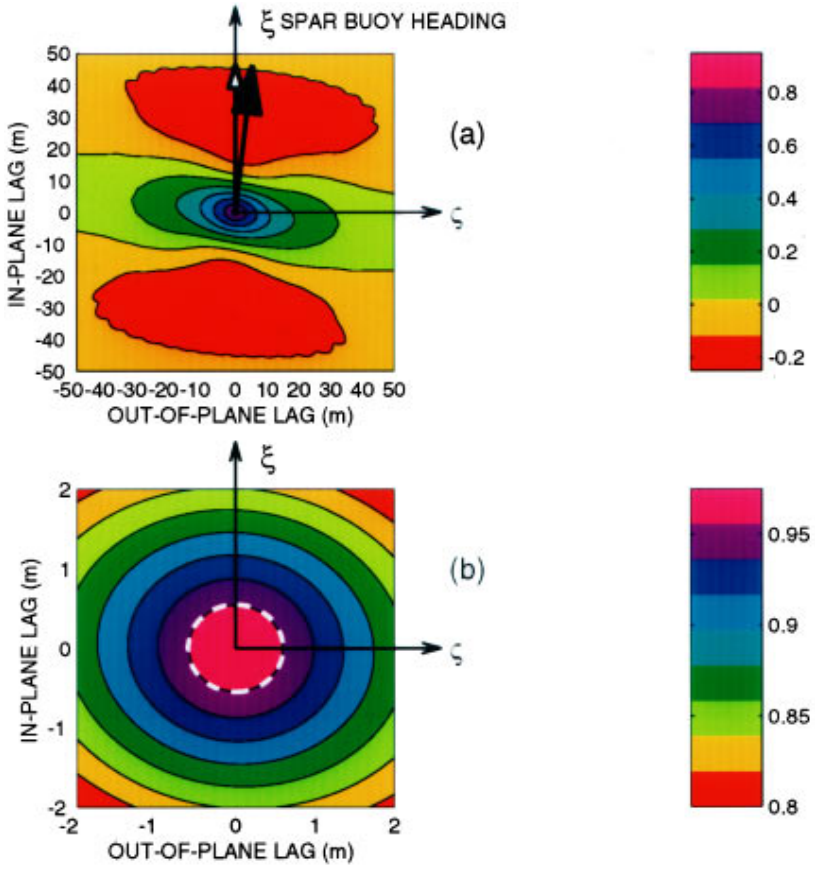


FIG. 5. Two-dimensional surface correlation function $C(\xi,\varsigma)$ for the wave conditions and orientation of the spar buoy during runs 2 and 3 (the conditions for run 4 produce a similar result); ξ corresponds to the in-plane axis (in line with the heading of the spar buoy) and ς to the out-of-plane axis. The magnitude of $C(\xi,\varsigma)$ is given by the color bars on the right. (a) Large-scale, low-resolution plot of $C(\xi,\varsigma)$ (not used in the scattering calculations) showing the mean wave direction (black arrowhead) and mean wind direction over previous hour (white arrowhead). (b) High-resolution plot of $C(\xi,\varsigma)$ (used in the scattering calculations) showing delineation of the effective correlation boundary (white, dashed line).

in Fig. 5(b). It is clear from the figure that for this run (and also runs 3 and 4) near-specular scattering is essentially independent of direction, as evidenced by the nearly circular effective correlation boundary of diameter 1.1 m. In other words, a rotation of $C(\xi,\varsigma)$ in order to study a different source-to-receiver direction would give the same estimate for the bistatic cross section. To be sure, the boundary is very slightly ellipsoidal, with a semi-minor axis oriented toward the vertical in the figure. This is expected, as the large-scale wave field was nearly unidirectional and oriented approximately 7 degrees to the right of the spar-buoy heading. The correlation function will always decay faster along this direction (perpendicular to the wave crests). This is best seen in low-resolution view of $C(\xi,\varsigma)$ [Fig. 5(a)], which shows the first minima of $C(\xi,\varsigma)$, located about ± 30 m from the origin. These minima are associated with a damped cosine component of $C(\xi,\varsigma)$ with a wavelength corresponding to the peak wave frequency of 0.16 Hz, or about 60 m. But because of the large value of χ, effective correlation is reached close to the origin of $C(\xi,\varsigma)$, where the contours of equal correlation are nearly circular owing to the increased directional spread of the shorter waves.

To test the conjecture that rotation of $C(\xi,\varsigma)$ has no significant effect on the bistatic cross section, we computed σ_r using the 2-D Kirchhoff integral [Eq. (1)] for various rotations of $C(\xi,\varsigma)$, and also using the 2-D isotropic version of the Kirchhoff integral [Eq. (4)], followed in each case by a conversion to the SSA(2) via Eq. (5). For the isotropic calculations, we used an equivalent low-pass version of $C(\rho)$, i.e., supported only by length scales $\geqslant 2\pi/K_{\text{sep}}$, which was based on a directionally averaged wave-number spectrum $F(K)$ (Appendix A). The results were essentially indistinguishable for the experimental conditions that existed during runs 2–4.

Next, we generated a $C(\rho)$ whose corresponding $F(K)$ now included wave numbers up to $K=500$ (essentially infinity), for use in calculating the bistatic cross section. For this $F(K)$ we used a wave spectral model developed by Plant[17] for the high-wave-number, or gravity-capillary, region, which was blended together with the long-wave components derived from the wave-buoy data (Appendix A). The relation between this $F(K)$ and $C(\xi,\varsigma)$ is such that they have equal s_L^2 (0.017), i.e.,

$$H^2\nabla^2 C(\xi,\varsigma)_{\text{origin}} = \int_0^{K_{\text{sep}}} K^3 F(K)\, dK \tag{14}$$

but the total mean-square slope represented in $F(K)$ and computed by integrating up to $K=500$ is 0.024.

Figure 6 summarizes these evaluations for the geometry and conditions that apply to run 2. Two cuts of model calculations of the bistatic cross section of the sea surface are shown. One [Fig. 6(a)] represents in-plane scattering such that the bistatic angle ϕ_b is always 0 degrees, and the other [Fig. 6(b)] represents out-of-plane scattering starting at the specular point (at range 32 m) where $\phi_b=0$ and moving perpendicularly 10 m to $\phi_b=31$ degrees. As a reminder, these calculations are based on either Eq. (1) or (4), with the results of each then being converted to the second-order small slope approximation SSA(2). A third result, also based

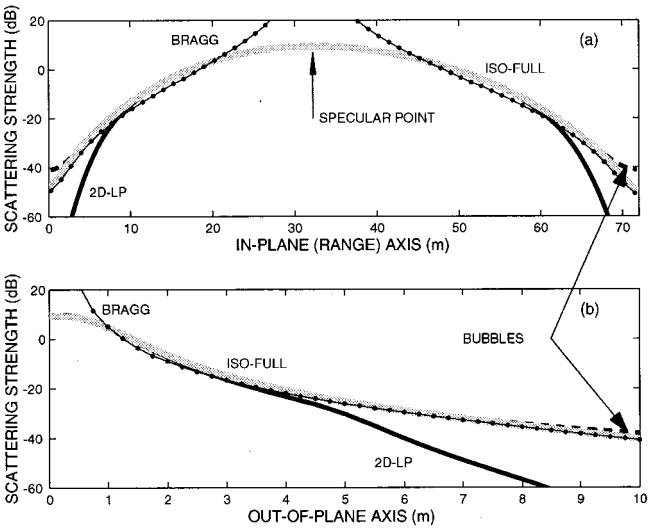


FIG. 6. Model estimates of the bistatic cross section σ as a function of (a) range along the in-plane axis, and (b) distance along the out-of-plane axis at range corresponding to the specular point (32 m) where $\phi_b = 0$. Dark, solid line (labeled 2D-LP) are results based on the 2-D, low-pass version of the correlation function. Light, solid line (labeled ISO-FULL) are results based on isotropic correlation function with full wave-number support, and the (mostly hidden) dashed line is the same computation with the effect of bubbles included. Dotted line (labeled BRAGG) are results based on the first-order perturbation approximation. The experimental conditions and scattering geometry are for run 2.

on Eq. (4), includes the contribution from near-surface bubbles. For this we use $\beta_I = 0.0012$ in Eqs. (12) and (13), a value estimated from surface backscattering measurements at low grazing angle taken during the course of the experiment.[21] These measurements were inverted for β_I using an approach described in Ref. 20, and we consider 0.0012 to be an upper bound for β_I associated with these conditions. Finally, for comparison, a fourth result shows σ_r computed with first-order perturbation theory alone [Eq. (7)]. To this we have not added the very small correction for tilt in Eq. (10), the results of which cannot be seen on the scale of these plots, nor the contribution from bubbles.

The first feature to note in Fig. 6 is the overlap within the near-specular region between the results of Eq. (1), which uses the 2-D correlation function incorporating wave directional information and wave numbers up to $K_{\text{sep}} = 31$ (dark, solid line labeled 2D-LP for low pass), and those of Eq. (4), which uses the isotropic correlation function incorporating wave numbers up to 500 (light, solid line labeled ISO-FULL for isotropic and inclusion of the full wave-number spectrum). The in-plane cut [Fig. 6(a)] shows the maximal scattering level at the exact specular point with a broad decay in level fore and aft of this point. Away from the specular region, there is a more rapid falloff in scattering level in the 2D-LP curve, e.g., at about 10 and 60 m in the in-plane cut shown in Fig. 6(a), and at about 4 m in the out-of-plane cut shown in Fig. 6(b). For the 2D-LP result, scattering associated with higher wave numbers is severely attenuated since only a large-scale, or low-pass, version of $C(\xi,\varsigma)$ is used in the scattering calculations. Note that, in principle, we could have generated a $C(\xi,\varsigma)$ with high-wave-number support, but the matrix size and integration task become unnecessarily burdensome, especially since wave directivity is becoming more isotropic with increasing wave number. (This property of waves eventually changes for very small capillary waves, whose wavelengths are a few millimeters and which respond rapidly to gusts of wind and then quickly succumb to viscosity.) Next, away from the near-specular region, the result of the SSA(2) calculations based on a $C(\rho)$ supported by all wave numbers (i.e., up to $K = 500$) merges nicely with the result of first-order perturbation theory (dotted line labeled BRAGG). As expected, the result of perturbation theory diverges from the SSA(2) calculations within the near-specular scattering region, because here the Bragg wave number corresponds to large-scale relief and first-order perturbation theory is no longer valid although improved accuracy is available from higher-order perturbation theory.[22,9] These results provide a revealing demonstration of how the SSA(2) accommodates the continuous transition from the near-specular scattering regime to the Bragg scattering regime when using realistic (i.e., power-law) ocean wave-number spectra.

Finally, the contribution from near-surface bubbles (mostly hidden dashed line) becomes significant only at relatively large bistatic angles ($\phi_b \gtrsim 30°$), at which point the bubble contribution roughly equals the Bragg contribution. Looking ahead, both the Bragg and the bubble-scattering contributions have essentially no impact on the spatial coherence measurements obtained during this experiment. Certainly, under less modest wind conditions these contributions would increase. (Modeling the effect of increasing bubble concentration on spatial coherence is discussed in Ref. 2.) We will show that the small-scale or Bragg contribution does play a role, however, in the time spreading observed in the ensemble-averaged intensity.

Figure 7(a) and (b) shows a plan view of model calculations of the bistatic cross section calculated for the conditions during experimental runs 2 and 3, which were taken sequentially but with different source depths. Here again we use Eq. (4) together with Eq. (5), and $C(\rho)$ incorporates wave numbers up to $K = 500$. The major effect of the scattering geometry is easily seen in the pattern of the scattering: the characteristic width of the pattern scales approximately as the ratio $\sin(\theta_g)/(1+\text{RD/SD})$, where SD and RD are the source and receiver depths (Table I).

The conditions during run 1 were quite different from those during runs 2–4, with the wind speed being down by a factor of 4 ($\lesssim$1 m/s) and H reduced about half (0.0786 m). Furthermore, the wave-field direction was mixed, with waves having a frequency less than about 0.35 Hz coming from 225 degrees and those having a frequency greater than 0.35 Hz coming from 20–50 degrees. The wind records show that the lower-frequency waves are associated with wind from the previous day, whereas the higher-frequency waves are associated with winds that existed during the previous 2 to 12 h. Thus run 1 represents an interesting case wherein the wind over the previous 0 to 2 h (coming from 100 to 130 degrees) has no relation to the surface roughness. Figure 8(a) shows a larger-scale picture of $C(\xi,\varsigma)$ corresponding to the conditions during run 1, with arrows indicating directions toward which the wind and wave fields were propagating, again ref-

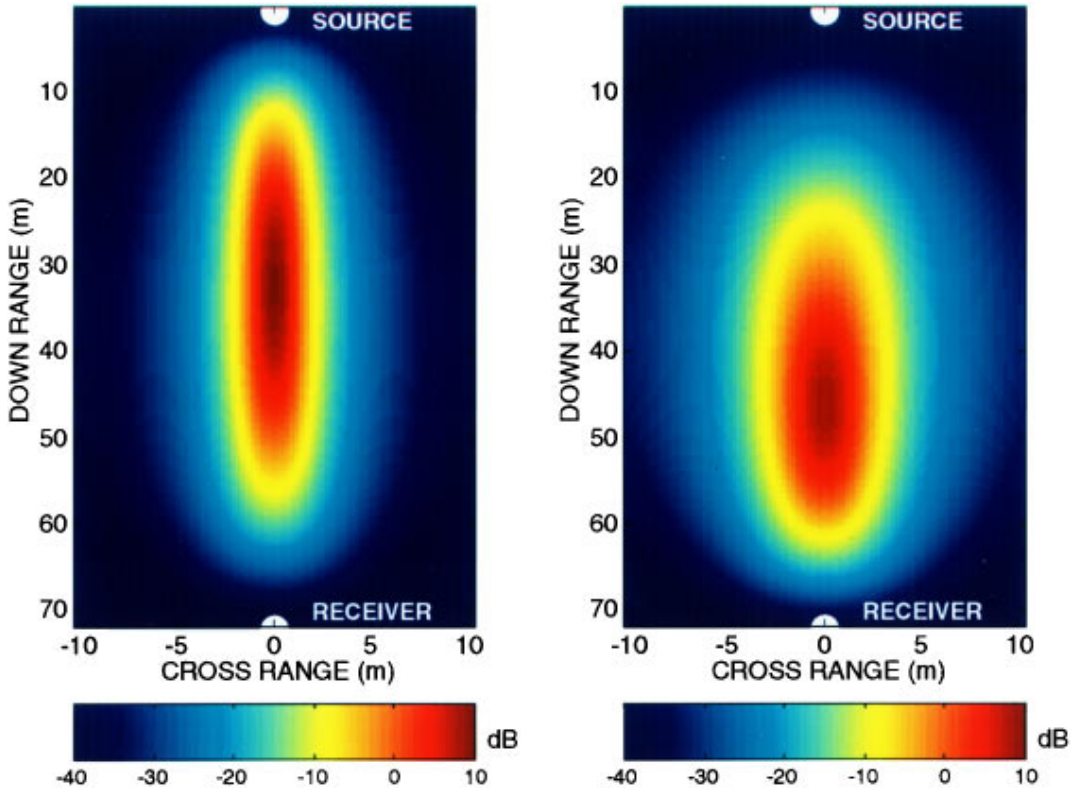


FIG. 7. Model estimates of the bistatic cross section σ as a function of position on the sea surface. (a) Experimental conditions and scattering geometry for run 2, with a source depth of 7.7 m and a receiver depth of 9.5 m, resulting in a nominal grazing angle of 14 degrees. (b) Experimental geometry for run 3 (same wave conditions as run 2), with a source depth of 17.7 m and a receiver depth of 9.5 m, resulting in a nominal grazing angle of 20 degrees. Note the aspect ratio for the axes, which exaggerates the cross-range dimension.

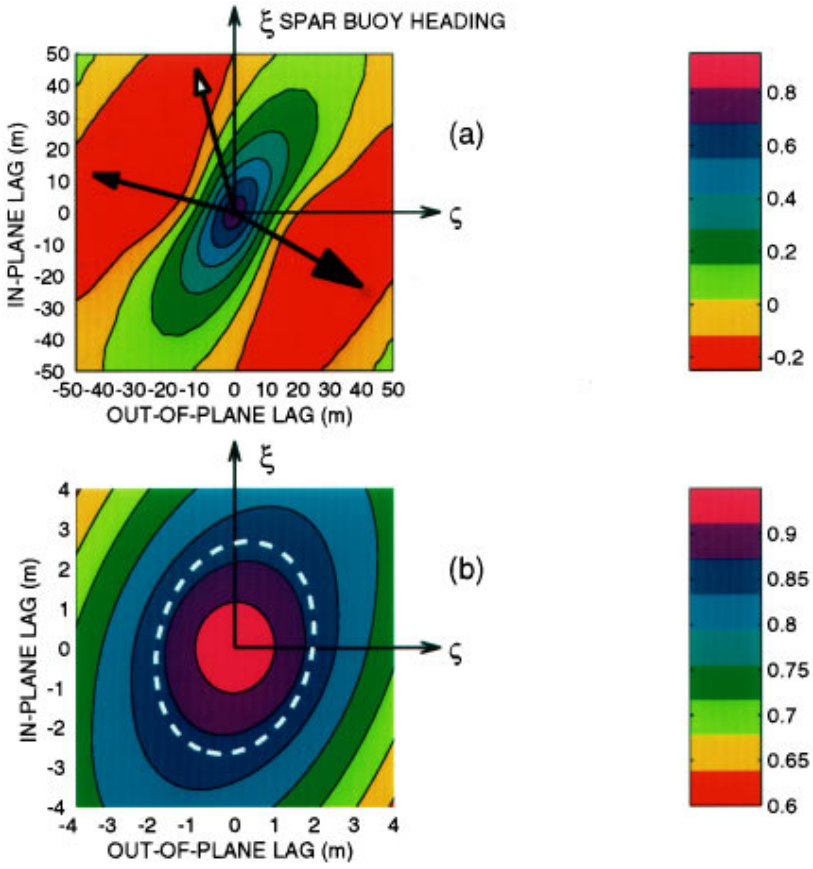


FIG. 8. Two-dimensional surface correlation function $C(\xi,\varsigma)$ for the wave conditions and orientation of the spar buoy during run 1; ξ corresponds to the in-plane axis (in line with the heading of the spar buoy) and ς to the out-of-plane axis. The magnitude of $C(\xi,\varsigma)$ is given by the color bars on the right. (a) Large-scale, low-resolution plot of $C(\xi,\varsigma)$ (not used in the scattering calculations) showing the mean wave direction for waves with frequency less than 0.3 Hz (large, black arrowhead), mean wave direction for waves with frequency greater than 0.3 Hz (small, black arrowhead), and mean wind direction over previous hour (white arrowhead). (b) High-resolution plot of $C(\xi,\varsigma)$ (used in the scattering calculations) showing delineation of the effective correlation boundary (white, dashed line).

erenced to the spar buoy heading. The dominant low-frequency wave field with a peak frequency of 0.16 Hz is traveling about 120 degrees to the right of the spar-buoy heading, and the higher-frequency waves are traveling in nearly the opposite direction, or 75 degrees to the left of the spar-buoy heading. Figure 8(b) shows the collection of points satisfying $e^{-\chi^2[1-C(\xi,\varsigma)]}=0.1$, which forms a slightly more elliptical boundary. The angle between the semiminor axis of this ellipse and the ξ axis of $C(\xi,\varsigma)$ is about 75 degrees [i.e., aligned with the arrow for higher-frequency waves in Fig. 8(a)].

Figure 9 shows expanded views of two cuts of model calculations of the bistatic cross section of the sea surface. One represents in-plane scattering such that the bistatic angle ϕ_b is always 0 [Fig. 9(a)], and the other represents out-of-plane scattering starting at the specular point where $\phi_b=0$ and moving perpendicular 1 m to $\phi_b=2.6°$ [Fig. 9(b)]. The solid line is the result when using the $C(\xi,\varsigma)$ shown in Fig. 8(b) (the expected orientation of the waves with respect to the spar buoy heading), and the dashed line is the result when using a $C(\xi,\varsigma)$ rotated 75 degrees. The rotated $C(\xi,\varsigma)$ produces an elliptical effective correlation boundary whose semiminor axis is aligned with the spar-buoy heading. The maximum difference between the two plots is 2 dB, and the qualitative difference is anticipated by the orientation of the respective ellipses: the true orientation (solid line) produces a slightly broader scattering pattern along the out-of-plane axis [Fig. 9(b)] because the distance between the origin of $C(\xi,\varsigma)$ and the effective correlation boundary measured along the out-of-plane (ς) axis is somewhat less than in the rotated case. The opposite relation exists for the in-plane (range) axis [Fig. 9(a)], and the two cases give the same result at the specular point (range 76 m), as expected from Eq. (4).

We anticipate that the directionality of the swell wave field has less impact than the directionality of the higher-frequency waves because the latter carry more of the wave-slope variance, which affects the scattering. But it is also clear that χ is a least one parameter controlling the effective correlation properties of the Kirchhoff integral Eq. (1). For run 1, $\chi=4.4$, and $C(\xi,\varsigma)$ must decay to about 0.88 to satisfy our criterion of the visibility function reaching 0.1. In contrast χ is $O(10)$ for the other three runs, and $C(\xi,\varsigma)$ need decay only to about 0.97. The decay of $C(\xi,\varsigma)$ is nearly isotropic close to the origin, i.e., approximating the equivalent $C(\rho)$, but becomes more anisotropic farther away from the origin where the 0.88 contour lies, because here slightly longer and more directive spatial lags become operative. This same effect can be observed in microwave measurements of sea surface backscattering. For example, measurements of the normalized radar cross section at 10 degrees (near-vertical) incidence angle show a dependence on azimuthal angle. This dependence is greatest at very low wind speeds and lessens (eventually disappearing) as the wind speed and waveheight variance increase.[19]

IV. RESULTS AND COMPARISON WITH FIELD MEASUREMENTS

Our primary acoustic observables are estimates of the magnitude of the horizontal spatial coherence, $|\hat{\Gamma}|$, and the time domain properties (e.g., time spreading and calibrated level) of the ensemble-averaged intensity, both of which are *integral* measures of the bistatic cross section. We use the method first described in Ref. 2 to link our model computations of the bistatic cross section to our measurements of horizontal coherence. Briefly, upon computing σ for a discretized 2-D sea surface (we use 0.25-m squares), a marginal density function for intensity versus horizontal arrival angle θ_h is generated; this function is then Fourier transformed to produce a model for the horizontal spatial coherence which we compare with our field measurements.

Figure 10 shows estimates of the magnitude of the horizontal coherence, $|\hat{\Gamma}|$, for the four data sets listed in Table I plotted against element spacing d normalized by wavelength, $kd/(2\pi)$. Error bars represent plus or minus one standard deviation at selected values of $|\hat{\Gamma}|$, based on[23]

$$\mathrm{Var}(|\hat{\Gamma}|)=\frac{(1-|\hat{\Gamma}|^2)^2}{2N}, \tag{15}$$

where N equals the number of pings used in the ensemble average (10 for runs 1 and 4 and 20 for runs 2 and 3). Each plot shows three model curves with complete overlap implied when only one can be seen. The black line is the final result for horizontal spatial coherence as derived from our model for the bistatic cross section. This calculation incorporates the true heading of the spar buoy, the directional wave field referenced to this heading, and its corresponding 2-D correlation function $C(\xi,\varsigma)$, for which spatial scales up to K_{sep} are included. Model/data agreement is quite satisfactory; for example, in all four data sets the model predictions lie fully within the estimated error bars of the data. The dashed line shows the model predictions for a hypothetical rotation of the scattering geometry. This rotation is 75 degrees for run 1, which aligns the semimajor axis of the ellipse drawn in Fig. 8(b) along the ς axis and therefore maximizes any anisotropic effects. The rotation is 90 degrees for runs 2–4, as there is no preferred orientation of the effective correlation boundary for these runs. Finally, the gray line is the model result when we used a directionally averaged wave-number spectrum $F(K)$ and corresponding $C(\rho)$ for computing σ. In this case, $C(\rho)$ is supported by length scales up to $K=500$. [Note: Upon repeating this calculation with a low-pass version of $F(K)$ based on wave numbers up to only K_{sep}, the result was precisely the same.] This result is anticipated from the curves shown in Fig. 6(a) and (b), which overlap in the near-specular region.

Referring now to Fig. 10(a), the correct spar-buoy heading with respect to the directional wave field produces a marginally better fit to the data (although all three model curves agree well with the data). The difference can be quantified by the characteristic normalized coherence length, $kd^{\star}/(2\pi)$, which we define as the number of wavelength separations needed for the model to decay to a value of $e^{-1/2}$. Using the correct heading gives a $kd^{\star}/(2\pi)$ of 20, whereas the 75-

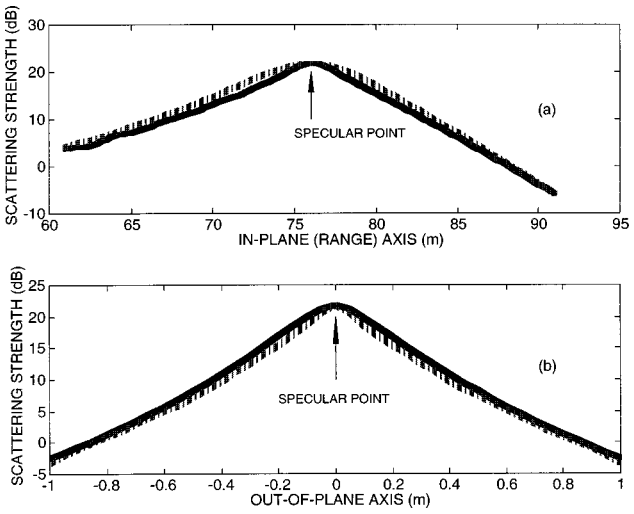


FIG. 9. Model estimates of the bistatic cross section σ as a function of (a) range along the in-plane axis, and (b) distance along the out-of-plane axis at range corresponding to the specular point (76 m). Experimental conditions and scattering geometry are for run 1. Solid line is the result of using the $C(\xi,\varsigma)$ shown in Fig. 8(b) (the true spar buoy heading); dashed line is the result of using a $C(\xi,\varsigma)$ rotated 75 degrees clockwise. This rotation aligns the semimajor axis of the ellipse shown in Fig. 8(b) along the ς axis.

degree rotation increases $kd^{\star}/(2\pi)$ to 23. This difference is anticipated upon inspection of the effective correlation boundary for $C(\xi,\varsigma)$ shown in Fig. 8(b): the two values of $kd^{\star}/(2\pi)$ scale approximately with the inverse of the distance between the origin of $C(\xi,\varsigma)$ and the effective correlation boundary as measured along the ς, or out-of-plane, lag axis. For runs 2–4 [Fig. 10(b)–(d)] the three model results (true heading, 90-degree rotation, and isotropic) are slightly less distinguishable in the important coherence range of $|\Gamma| \lesssim 0.6$, which is again anticipated from the more circular effective correlation boundary shown in Fig. 5(b).

We use a Gaussian function,

$$\Gamma(kd) = \exp[-(kd\sigma_{\theta_h})^2/2], \tag{16}$$

as a convenient single-parameter model with which to summarize the measured data, and interpret $1/(kd^{\star})$ as being equal to the rms horizontal angular spread σ_{θ_h} resulting from the forward scattering. Values of $kd^{\star}/(2\pi)$ and σ_{θ_h} for all four data sets are summarized in Table I; since the measured coherence goes below $e^{-1/2}$ for only one case (run 3), these values are based on the model curves (thicker black line) in each case. For run 1, the net effect of the 75-degree rotation is to reduce σ_{θ_h} by about 0.5 degrees. Note again that the data shown in Fig. 10(b) and (c) (runs 2 and 3 from Table I) were taken sequentially but with different source depths and illustrate the major effect that array acquisition geometry has on spatial coherence. Changing the source depth from 7.7 to 17.2 m increases σ_{θ_h} by a factor somewhat greater than 2, a factor that approximately scales with the ratio of the values of $\sin\theta_g/(1+\mathrm{RD/SD})$ for the two geometries.

It is clear that measurements of horizontal spatial coherence are a useful surrogate for measurements of the bistatic cross section, σ, close to the specular region, where spatial gradients in σ are very sharp. A practical implication of the spatial coherence estimates and their reduction to σ_{θ_h} is that a bound can be placed on improvements in the directional resolution ($\approx\lambda/L_A$) of surface bounce paths achieved by increasing the array length L_A. Since the geometry and medium (in this case the surface boundary) impose a directional resolution of approximately $2\sigma_{\theta_h}$ the *effective* array length is bounded by $L_A < \lambda/2\sigma_{\theta_h}$. For example, the combination of surface conditions and array acquisition geometry that existed during run 3 suggests that the effective array length at 30 kHz is bounded by about 75 cm.

We now return to Figs. 2 and 3 and the model curves plotted over the ensemble-averaged intensity data in the surface-bounce region [one in Fig. 2(b) and two in Fig. 3(b)]. These curves are the result of convolving a model for the intensity impulse response $I_{\mathrm{imp}}(t)$ with a 1-ms square wave (representing the 1-ms transmit pulse used in the experiment). The model for intensity impulse response is, in turn, generated by summing over the discretized sea surface such that

$$I_{\mathrm{imp}}(t) = I_0 \sum_n \frac{\sigma_n 10^{\alpha_w (R_{i_n}+R_{s_n})/10} dA}{R_{i_n}^2 R_{s_n}^2}, \tag{17}$$

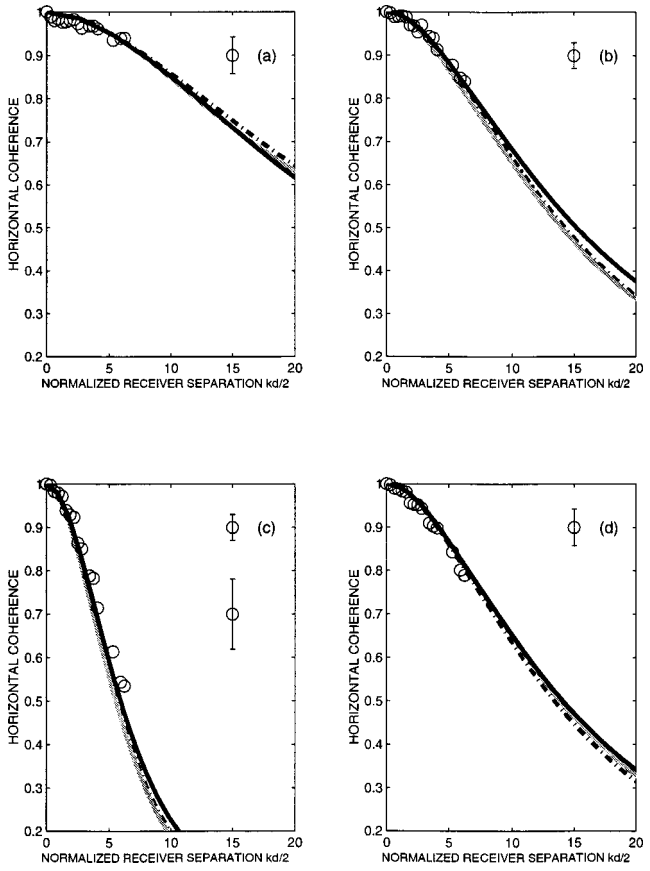


FIG. 10. Horizontal coherence versus normalized receiver separation $kd/(2\pi)$ with plots (a)–(c) corresponding to runs 1–4. Error bars represent plus or minus one standard deviation for coherence magnitudes of 0.9 and 0.7. The thick, black line is the final result derived from the model for the bistatic cross section when using the true heading of the spar buoy, the directional wave field referenced to this heading, and its corresponding 2-D correlation function $C(\xi,\varsigma)$. The dashed line is the result for a hypothetical rotation of the scattering geometry. This rotation is 75 degrees clockwise for run 1 [(a)] and 90 degrees clockwise for runs 2–4 [(b)–(d)]. The gray line is the result of using an isotropic correlation function $C(\rho)$ in the computations, which is equal to the Bessel transform of a directionally averaged wave-number spectrum $F(K)$.

where the area dA is 0.0625 m^2 for the nth scattering patch with bistatic cross section of σ_n. The sum of the incident (R_{i_n}) and scattered (R_{s_n}) slant ranges is mapped to time t via a constant sound speed (an excellent approximation for these experimental conditions). The effects of the transmit beam pattern can be safely ignored, and a small correction is made for the nearly omnidirectional receive beam pattern. We use 0.005 dB/m for the sea-water attenuation, α_w, and the source level, I_0, was 176 dB *re* 1 μPa at 1 m for this experiment.

In Fig. 2(b), corresponding to run 1, the model curve matches the data very well over the maximum steady-state phase, plus the next approximately 23 dB of decay in the data time series. The level of the data then abruptly changes owing to the onset of the first bottom arrival. Two model curves were actually computed, one based on scattering calculations incorporating wave numbers up to a maximum of $K=500$, and the other incorporating wave numbers up to a maximum $K_{\text{sep}}=31$. There is absolutely no difference in these results [i.e., the two curves overlay on the scale of Fig. 2(b)], suggesting that Bragg-type bistatic scattering from small-scale, wind-generated waves is insignificant. We would expect this given the calm, low-wind, and diminished wave conditions that existed during run 1 (see Table I), and given the reduced time spreading seen in the surface arrival data.

In Fig. 3(b), corresponding to run 2, there is now a clear difference between the two model curves incorporating wave numbers up to a maximum of $K=500$ (solid curve) and up to a maximum $K_{\text{sep}}=31$ (dotted curve). Both curves match the level of the data well during the maximum steady-state phase, plus the next 20 dB of temporal decay. But only the calculation that incorporates higher wave numbers associated with Bragg-type scattering matches the additional 10 dB of decay that is clearly seen in the data. The level of the data then abruptly increases again, owing to the onset of the bottom-bounce arrival.

We find that impulse response $I_{\text{imp}}(t)$ goes approximately as $e^{-t/L}$ for $t>L$, where L is a characteristic time scale which best fits the response. This value is about 0.03 ms for run 1, and about 0.21 ms for run 2. It is also about 0.21 ms for run 4, for which the geometry and conditions were similar to those of run 2, but L increases to about 0.31 ms for run 3. The inverse, $1/L$, is an approximate measure of the frequency coherence bandwidth.[24] In our case the frequency coherence decreases to roughly $e^{-1/2}$ for frequency separations greater than about $1/2L$. For run 2, $1/2L$ is $\sim$2.4 kHz, and decreases to $\sim$1.6 kHz for run 3 owing to the simple change in source depth, while for the relatively calm conditions of run 1 the bandwidth increases to about $\sim$17 kHz.

V. SUMMARY

Measurements have been presented of both the spatial coherence and ensemble-averaged intensity of sound that has been forward scattered from the sea surface, along with a model for the sea surface bistatic-scattering cross section, σ. The measurements represent integral properties of bistatic scattering, and the model for σ was used to generate these same properties for comparison with the field data. The measurements were made at 30 kHz in a shallow-water channel of depth 25.6 m. The coherence was measured across a line array oriented transverse to the direction of propagation, giving an estimate of the horizontal coherence or, equivalently, the horizontal angular spread as related by Fourier transform.

Key environmental parameters were monitored by a directional wave buoy operating within 100 m of the experiment, a ship-borne anemometer (co-located with the receiving line array), and CTD casts (conducted from the research vessel). The conditions were mild, with wind speeds of 1 to 4 m/s and rms wave heights of 8 to 16 cm. Nevertheless, we show that changes within these ranges influenced the measurements to a degree predicted by our model.

The bistatic cross section, σ, was divided into two components: σ_r, associated with scattering from the rough air/sea interface, and σ_b, associated with scattering from near-surface bubbles. We used the second-order small slope approximation, SSA(2), to compute σ_r. The primary advantage of the small slope approximation was the resulting smooth transition in σ_r over a broad range of scattering angles.

The Kirchhoff integral remains as the computational basis for SSA(2), with the corresponding integrand controlled by the surface roughness parameter, χ, and the two-dimensional surface-wave-height autocorrelation function, $C(\xi,\varsigma)$. Estimates of $C(\xi,\varsigma)$ were generated by combining measurements of the directional wave spectrum made with a wave buoy and a model for the directional spreading of waves whose temporal frequency exceeded the maximum frequency measured by the buoy. Key properties of $C(\xi,\varsigma)$ in the context of bistatic scattering were expressed in terms of an effective correlation boundary, the point at which $e^{-\chi^2[1-C(\xi,\varsigma)]}$ dropped below 0.1. This boundary was nearly circular for the combination of environmental conditions and scattering geometry existing during three of the four experimental runs, for which $\chi\gtrsim 10$. Thus it was anticipated that for these experimental runs, the exact orientation of the source-to-receiver direction *vis-à-vis* the surface-wave-field direction was largely irrelevant, a finding that was confirmed by simulation. In the experimental run during which both the wind speed and wave height were the smallest, χ equaled 4.4, and the effective correlation boundary in this case formed an approximate ellipse. In this case, the simulation demonstrated that the orientation of the source-to-receiver direction *vis-à-vis* the surface wave field direction would have made a difference in forward scattering, although the effect still was quite small.

The analysis of the effective correlation properties of $C(\xi,\varsigma)$ suggested that an isotropic correlation function, $C(\rho)$, based on the directionally averaged wave-number spectrum, $F(K)$, would be equally effective in the scattering calculations. In fact, model/data agreement for horizontal spatial coherence was quite good for all four experimental runs studied, regardless of whether $C(\xi,\varsigma)$ or $C(\rho)$ was used in the scattering calculations.

To generate $C(\xi,\varsigma)$, we incorporated surface wave numbers up to $K_{\text{sep}}=0.25k$ (equal to 31 at our frequency of

30 kHz), and we viewed this as a large-scale, or low-pass, version of the correlation function. But for $C(\rho)$ we incorporated wave numbers up to $K=500$ by using a wave spectral model[17] for the high-wave-number, or gravity-capillary, region. However, repeating these calculations with a low-pass version of $C(\rho)$ (incorporating wave numbers only up to $K=31$) gave exactly the same results for horizontal spatial coherence. This confirms that spatial coherence is sensitive to bistatic scattering only from within the near-specular angular region, the properties of which are established by the large-scale components of the sea-surface roughness spectrum. Scattering associated with near-surface bubbles also had no effect on the model results; however, the conditions were relatively mild, and we expect that bubble scattering plays more of a role at higher wind speeds.

In contrast, we showed that temporal properties of ensemble-average intensity were sensitive to the Bragg scattering contribution. For example, fully 30 dB of decay in the measured time series of forward scattering was well matched by a simulated time series that included the Bragg scattering contribution from the high-wave-number, or gravity-capillary, region.

Finally, use of the small slope approximation effectively eliminates the need to partition the wave-number spectrum into, for example, the large-scale wave-number region defined by $K \leqslant K_{\text{sep}}$ and the remaining small-scale wave-number region. The small slope approximation accurately incorporates all wave numbers, producing a smooth transition in σ_r over a broad range of scattering angles. The partitioning idea remains useful, however, for computing a mean-square slope of the large-scale waves (s_L^2) in order to correct for the tilting of the large-scale surface upon which the smaller-scale (Bragg) waves propagate. An expression for tilt modulation is derived in Appendix B. The effects of tilt modulation were not readily observable in this experiment, however, owing to the very small values of s_L^2.

ACKNOWLEDGMENTS

This work was funded by the Office of Naval Research Code 321 Ocean Acoustics, via Contract No. N00039-91-C-0072 and Grant No. N00014-96-1-0325.

APPENDIX A: SURFACE WAVE SPECTRAL MEASUREMENTS AND MODELING, AND CONSTRUCTION OF SURFACE WAVE CORRELATION FUNCTIONS

In this appendix we describe the directional wave measurements made using a 1-m-diameter Datawell directional wave buoy (positioned within 100 m of the acoustic measurement site) and the method used to convert these data to an estimate of the spatial correlation function of sea-surface roughness for use in the scattering calculations. The wave buoy generated estimates of the surface wave height frequency spectrum, $S(\omega)$, its mean direction, $\overline{\phi}(\omega)$, and the rms directional spread about the mean direction, $\sigma_{\overline{\phi}(\omega)}$, all a function of frequency $\omega=2\pi f$. These data were telemetered back to the research vessel continually every hour.

The maximum resolvable frequency for the wave buoy is 0.58 Hz; beyond this frequency we extend the rear face of $S(\omega)$ up to a frequency of 2.8 Hz. This frequency is equivalent (through the dispersion relation $\omega^2=gK$, where g is a gravitional constant) to the separation wave number, K_{sep}, chosen to equal 0.25 times the acoustic wave number, k. We set $K_{\text{sep}}=31$ for our acoustic frequency of 30 kHz. Typically, frequency extension is achieved using a power law of the form $S(\omega)\sim\omega^{-n}$ for the spectral roll-off, where n is 4 to 5. Banner[25] provides clarification on the important issue of spectral decay for the rear face of the frequency spectrum; the upshot is that this decay must depend on AK_p, defined as the integral of $S(\omega)$ within the segment 0.5 to 1.5 times the peak frequency (equal to A), where K_p is the peak surface wave number. If this integral measure of $S(\omega)$ is sufficiently large, then the net Doppler-shifting influence of these waves (due to the orbital motion) will progressively lessen the magnitude of the roll-off exponent n from its asymptotic value of 5. Banner parametrizes the effect of the dominant waves in terms of AK_p; in our case, this parameter is quite small, being <0.01 for all our spectra, and thus we would expect the asymptotic dependence of ω^{-5} for the high-frequency gravity waves. However, at 0.58 Hz, the frequency roll-off of the wave-buoy data is closer to ω^{-4}. Therefore we argue that a smooth transition in roll-off starting at ω^{-4} and ending at ω^{-5} is more suitable for this data set. Accordingly, for the rear face of $S(\omega)$ we use a roll-off that goes as $\omega^{-n(f)}$, where the exponent of the frequency dependence changes linearly from -4 to -5 over the frequency range 0.58 to 2.8 Hz in a manner described by

$$n(f)=0.4504f+3.7390. \tag{A1}$$

The directional information from the wave buoy is used to construct a 2-D spatial correlation function, $C(\xi,\varsigma)$, which is, in turn, used in Eq. (1). For this task we rely on a model for the directional spreading of waves given by Donelan *et al.*,[26] hereafter referred to as DHH, plus a small modification to this model given by Banner. The directional frequency spectrum, defined as $S(\omega)D(\phi;\omega)$, is generated by combining the mean direction versus frequency, $\overline{\phi}(\omega)$, derived from the wave buoy and the DHH directional spreading function,

$$D(\phi;\omega)=\frac{b}{2}\operatorname{sech}^2[b(\phi-\overline{\phi}(\omega)]. \tag{A2}$$

The spreading parameter b is a function of the ratio ω/ω_p, where ω_p is the peak frequency of the spectrum, and is prescribed by

$$b=\begin{cases} 1.24, & \omega/\omega_p<0.56, \\ 2.61(\omega/\omega_p)^{1.3}, & 0.56<\omega/\omega_p<0.95, \\ 2.28(\omega/\omega_p)^{-1.3}, & 0.95<\omega/\omega_p<1.6, \\ 10^{-0.4+0.8393\exp[\log_e[(\omega/\omega_p)^2]]}, & \omega/\omega_p>1.6. \end{cases} \tag{A3}$$

On the basis of more recent observations, Banner determined that wave directional spreading should increase beyond the high-frequency limiting value that DHH originally prescribed ($b=1.24$ for $\omega/\omega_p>1.6$), and thus we use Banner's modification for $\omega/\omega_p>1.6$ in Eq. (A3). Figure A1 shows an

example of the frequency spectrum and directional spread estimated from the wave buoy, plotted along with an equivalent directional spread derived from the DHH model in Eqs. (A2) and (A3). We define the equivalent directional spread as equal to one-half the width of Eq. (A2) (normalized to a maximum value of unity) between $e^{-1/2}$ points, or an equivalent Gaussian standard deviation. The frequency spectrum is the one associated with runs 2 and 3, during which the wind speed was 4 m/s.

The equivalent spread values derived from the DHH model plus Banner's modification match well with the spread estimates obtained from the wave buoy, which go up to 0.58 Hz; however, the model can be extended to higher frequencies to meet our needs. To avoid unnecessary noise in subsequent calculations, we use the DHH spreading model over the entire frequency range (up to 2.8 Hz) and use the buoy data for the mean direction, $\bar{\phi}(\omega)$, for frequencies up to 0.58 Hz. The mean direction is assumed to be constant for the frequency extrapolation up to 2.8 Hz, with the constant equal to the average of the mean directions corresponding to the last four frequency bins (between 0.55 and 0.58 Hz).

The normalized correlation function, $C(\xi,\varsigma)$, is then computed via numerical implementation of

$$C(\xi,\varsigma)=\frac{1}{H^2}\int_{-\pi}^{\pi}d\phi\int_{0}^{5.6\pi}d\omega\,S(\omega)D(\phi;\omega)\cos[(\omega^2/g)\times(\xi\cos\phi+\varsigma\sin\phi)]. \tag{A4}$$

Note the implied equivalence between surface wave number K and ω^2/g [also, shallow water effects are insignificant at depth $\approx$26 m with $\tanh(K_p\cdot\text{depth})\approx 1$]. The effects of surface tension are significant only for wavelengths $\lesssim$7 cm (or wave temporal frequencies greater than 4.7 Hz), and complications owing to Doppler shifting are absent (refer to the foregoing remarks on Banner's work). This justifies our using the gravity wave dispersion relation for frequencies up to 2.8 Hz in computing $C(\xi,\varsigma)$, and we include in the construction of $C(\xi,\varsigma)$ length scales down to size $L_{\min}=2\pi/K_{\text{sep}}$.

We also compute an isotropic correlation function, $C(\rho)$, for use in Eq. (4), via the Bessel transform relation

$$C(\rho)=\frac{1}{H^2}\int_{0}^{\infty}KJ_0(K\rho)F(K)dK, \tag{A5}$$

where $F(K)$ is the directionally averaged wave-number spectrum. For $F(K)$ we use an amalgamation of a low-wave-number segment of the wave-number spectrum derived from the wave buoy data and a high-wave-number segment computed using a surface wave vector model derived by Plant.[17] The Plant model requires both wind speed and fetch as input and produces a directional wave-number spectrum with a spectral peak that depends on fetch. It, too, now uses the DHH and Banner spreading functions. We take a directional average of Plant's model result and use it for $K>10$ in the amalgamated spectrum. Wave numbers up to $K=500$ are included in $F(K)$, at which point $F(K)$ is rapidly but smoothly attenuated with a Gaussian tapering function; thus we include in the corresponding $C(\rho)$ length scales down to size $L_{\min}=2\pi/500$. In Sec. IV, we also mention a low-pass version of $C(\rho)$, which we compute by using an $F(K)$ that also is rapidly and smoothly attenuated but for wave numbers greater than K_{sep}.

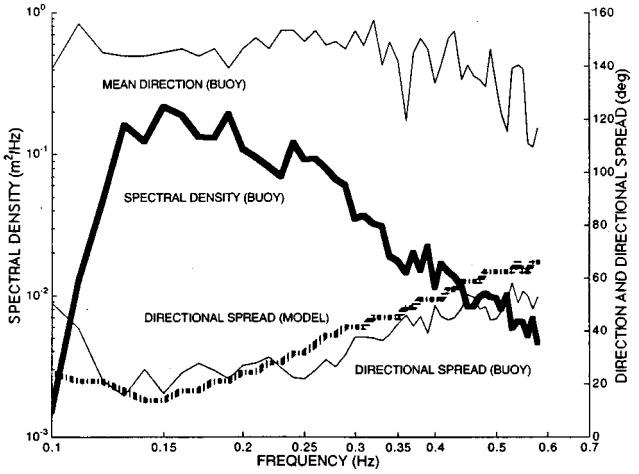


FIG. A1. Frequency spectrum, mean direction, and directional spread, estimated by the wave buoy, for the conditions in effect during runs 2 and 3; rms wave height, H, is 0.16 cm, and the peak frequency is approximately 0.16 Hz. The dashed line is an equivalent directional spread derived from the DHH model.

Figure A2 shows the resulting $F(K)$ spectrum, again associated with runs 2 and 3. The thin, black line is the original buoy data plus the frequency extension to 2.8 Hz based on Eq. (A1), expressed in the wave-number domain. The thick, gray line is the resulting amalgamation between it and the Plant model, where the model portion was generated using a wind speed equal to 4 m/s and a fetch of 20 km. [Note that the fetch input is immaterial here, since $F(K)$ is largely immune to fetch for K greater than about 10 and when the fetch is greater than about 1 km.] The two segments overlap within an equilibrium range of the spectrum, roughly between $K=1$ and $K=10$; within this range, the

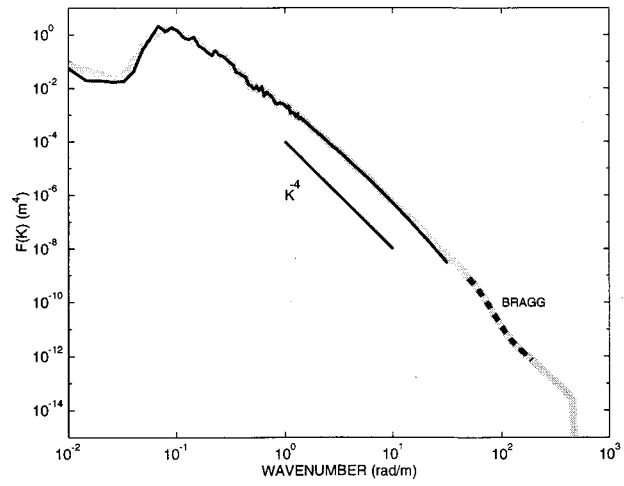


FIG. A2. Directionally averaged wave-number spectrum $F(K)$ for runs 2 and 3. The thin, black line is the original buoy data plus frequency extension to 2.8 Hz based on Eq. (A1); the thick, gray line is the resulting amalgamation between it and the Plant model. Overlap occurs within the equilibrium range of the spectrum, roughly in the range K equals 1 to 10. The thick, dashed segment corresponds to the approximate range of Bragg wave numbers that applies to the bistatic scattering conditions and scattering geometries for runs 2 and 3.

spectrum is least sensitive to changes in wind speed and tends to follow a wave-number spectral dependence of about K^{-4}.[25]

The thick, dashed segment within $F(K)$ corresponds to the approximate range of Bragg wave numbers that applies to the bistatic scattering conditions and geometries of runs 2 and 3. In other words, it is only within this range of wave numbers that perturbation theory gave a result comparable to that of the small slope approximation (and therefore was considered accurate), as was illustrated in Fig. 6(a) and (b).

Finally, we remark here that all three waveheight frequency spectra used in this study (and many others obtained over the course of the experiment) have a peak frequency near 0.16 Hz, which is too low to be associated with the modest wind speeds involved and is instead associated with aging wind waves originating from a southerly direction (Gulf of Mexico). The wind speed associated with the spectrum shown in Figs. A1 and A2 was 4 m/s, but during the previous 12 h wind speed had averaged closer to 6 m/s. A study by Nguy,[27] on the entire set of wave data gathered during the experiment, suggests that this kind of prevailing wind condition—winds on the decrease along with reduced wind forcing—tends to favor nonlinear wave–wave interaction as the controlling feature in spectral evolution. Nonlinear wave–wave interaction will cause an evolution of the spectral peak toward lower frequencies[28] and has also been shown to have a stabilizing effect on the spectral shape, driving the wave spectrum to unimodal form.[29] Interestingly, the DHH spreading model which, strictly speaking, describes spreading for a wind-generated deep-water directional spectrum, works quite well in the mixed swell and wind seas represented here, all of which were of unimodel form. If we consider the wave age parameter, c_p/U, where c_p is the phase speed associated with the spectral peak and U is wind speed, this value is about 2.5 for runs 2–4 and about 10 for run 1. Putting this into perspective, a so-called fully developed spectrum, such as the Pierson–Moskowitz spectrum, is characterized by a wave age parameter of 1.2. The spectra measured in this experiment, then, can be viewed as *over* developed and represent clearly more complicated mixed seas, for which the wind speed has little to do with either the total rms waveheight, H, or the peak frequency. However, if one instead examines properties, such as rms waveheight and rms slope, associated with higher frequency bins, say between 0.45 and the maximum 0.58 Hz, a much better correlation with local wind condition is found.[27]

APPENDIX B: CORRECTION FOR TILT MODULATION IN THE BISTATIC CROSS SECTION

In this appendix we derive a correction term that accounts for the modulation of θ_i and θ_s owing to the tilt imposed by the large-scale surface. Consider a patch of sea surface, or facet, of dimension L_F which is large in comparison with the acoustic wavelength, λ, yet small enough to be considered a plane surface insofar as its normal $\mathbf{n}$ is approximately constant. With Bragg scattering theory also requiring that the rms relief on the facet be small in comparison with λ, the above requirements are reasonably satisfied by $L_F = 4$ to $5\,\lambda$ (cf. Refs. 7, 17, and 18).

The sea surface elevation is $\eta(x,y)$, and the pitch (say $d\eta/dx$) and roll ($d\eta/dy$) of the surface patch about the mean x,y plane compose a tilt angle $\delta = \sqrt{(d\eta/dx)^2+(d\eta/dy)^2}$. The two slopes $d\eta/dx$ and $d\eta/dy$ have been shown to be well described by a Gaussian distribution[8] and, if we assume further that their variances are each approximately equal to $s_L^2/2$, then δ is a Rayleigh-distributed random variable.

The projection of $\mathbf{n}$ onto the x,y plane forms the angle ψ with respect to some arbitrary reference axis. (It is helpful to assume, without loss of generality, that this reference axis is the x axis and lies in the same vertical plane as $\mathbf{k}_i$.) We assume that ψ is distributed uniformly between 0 and 2π and is independent of δ.

The effect of tilt modulation on bistatic scattering is then given by the expectation

$$\langle \sin(\theta_i+\Delta_i)^2 \sin(\theta_s-\Delta_s)^2 \rangle, \tag{B1}$$

where $\Delta_i = \delta\cos\psi$ and $\Delta_s = \delta\cos(\psi-\phi_b)$. To carry out the expectation we expand and retain terms of $O(\Delta_i^2,\Delta_s^2)$. Equation (10) follows upon noting that $E(\Delta_{i,s})=0$, $E(\Delta_{i,s}^2) = s_L^2/2$, and $E(\Delta_i\Delta_s) = s_L^2\cos(\phi_b)$.

[1] A. G. Voronovich, *Wave Scattering from Rough Surfaces* (Springer-Verlag, New York, 1993).

[2] P. H. Dahl, ''On the spatial coherence and angular spreading of sound forward scattered from the sea surface: Measurements and interpretive model,'' J. Acoust. Soc. Am. **100**, 748–758 (1996).

[3] D. R. Jackson, K. B. Briggs, K. L. Williams, and M. D. Richardson, ''Tests of models for high-frequency seafloor backscatter,'' IEEE J. Ocean Eng. **21**, 458–470 (1996).

[4] P. H. Dahl and W. L. J. Fox, ''Measurement and interpretation of angular spreading from multiple boundary interactions in a shallow water channel,'' in *High Frequency Acoustics in Shallow Water*, edited by N. G. Pace, E. Pouliquen, O. Bergem, and A. P. Lyons (SACLANT Undersea Research Centre, La Spezia, Italy, 1997), pp. 107–114.

[5] J. W. Wright, ''A new model for sea clutter,'' IEEE Trans. Antennas Propag. **AP-14**, 217–223 (1968).

[6] S. T. McDaniel and A. D. Gorman, ''An examination of the composite-roughness scattering model,'' J. Acoust. Soc. Am. **73**, 1476–1486 (1983).

[7] K. Hasselmann, R. K. Raney, W. J. Plant, W. Alpers, R. A. Shuchman, D. R. Lyzenga, C. L. Rufenach, and M. J. Tucker, ''Theory of synthetic aperture radar ocean imaging: A MARSEN view,'' J. Geophys. Res. **90**, 4659–4686 (1985).

[8] W. J. Plant, ''Bragg scattering of electromagnetic waves from the air/sea interface,'' in *Surface Waves and Fluxes*, edited by G. L. Geernaert and W. J. Plant (Kluwer Academic, Norwell, MA, 1990), Vol. II.

[9] E. I. Thorsos, ''Acoustic scattering from a Pierson-Moskowitz sea surface,'' J. Acoust. Soc. Am. **88**, 335–349 (1990).

[10] S. T. McDaniel, ''Diffractive corrections to the high-frequency Kirchhoff approximation,'' J. Acoust. Soc. Am. **79**, 952–957 (1986).

[11] R. Dashen, F. S. Henyey, and D. Wurmser, ''Calculations of acoustic scattering from the ocean surface,'' J. Acoust. Soc. Am. **88**, 310–323 (1990).

[12] E. I. Thorsos and S. L. Broschat, ''An investigation of the small slope approximation for scattering from rough surfaces. Part I. Theory,'' J. Acoust. Soc. Am. **97**, 2082–2093 (1995).

[13] S. L. Broschat and E. I. Thorsos, ''An investigation of the small slope approximation for scattering from rough surfaces. Part II. Numerical studies,'' J. Acoust. Soc. Am. **101**, 2615–2625 (1997).

[14] E. I. Thorsos and S. L. Broschat, ''The lowest-order small slope approximation for rough surface scattering,'' in Proceedings 16th International Congress on Acoustics and 135th Meeting Acoustical Society of America IV, 3015–3016, 1998.

[15] P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces* (Pergamon, Oxford, 1963).

[16] A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, New York, 1978), Vol. II.

[17] W. J. Plant, ''A two-scale model of short wind-generated waves and scatterometry,'' J. Geophys. Res. **91**, 10735–10749 (1986).
[18] S. T. McDaniel and D. F. McCammon, ''Composite-roughness theory applied to scattering from fetch-limited seas,'' J. Acoust. Soc. Am. **82**, 1712–1719 (1987).
[19] V. Hesany, W. J. Plant, and W. C. Keller, ''The normalized radar cross section of the sea surface at 10° incidence,'' J. Geophys. Res. (accepted).
[20] P. H. Dahl, W. J. Plant, B. Nützel, A. Schmidt, H. Herwig, and E. A. Terray, ''Simultaneous acoustic and microwave backscattering from the sea surface,'' J. Acoust. Soc. Am. **101**, 2583–2595 (1997).
[21] P. H. Dahl, ''High frequency surface backscattering: A 24-hr time series of data showing transition between Bragg and resonant bubble scattering,'' J. Acoust. Soc. Am. **100**, 2840 (1996).
[22] E. I. Thorsos and D. R. Jackson, ''The validity of the perturbation approximation for rough surface scattering using a Gaussian roughness spectrum,'' J. Acoust. Soc. Am. **86**, 261–277 (1989).
[23] G. C. Carter, C. H. Knapp, and A. H. Nuttall, ''Statistics of the estimate of the magnitude-coherence function,'' IEEE Trans. Audio Electroacoust. **21**, 388–389 (1973).
[24] R. S. Kennedy, *Fading and Dispersive Channels* (Wiley-Interscience, New York, 1969).
[25] M. L. Banner, ''Equilibrium spectra of wind waves,'' J. Phys. Oceanogr. **20**, 966–984 (1990).
[26] M. A. Donelan, J. Hamilton, and W. H. Hui, ''Directional spectra of wind generated waves,'' Philos. Trans. R. Soc. London, Ser. A **315**, 509–562 (1985).
[27] D. K. Nguy, ''Observations on the Directional Development of Wind-Waves in Mixed Seas,'' M.S. thesis, University of Washington, Department of Mechanical Engineering, Seattle, March 1998.
[28] K. Hasselmann, T. P. Barnett, E. Bouws, H. Carlson, D. E. Cartwright, K. Enke, J. A. Ewing, H. Gienapp, D. E. Hasselmann, P. Kruseman, A. Meerburg, P. Müller, D. J. Olbers, K. Richter, W. Sell, and H. Walden, ''Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP),'' Dt. hydrogr Z. A **8**, (Suppl)(12) (1973).
[29] D. Masson, ''On the nonlinear coupling between swell and wind waves,'' J. Phys. Oceanogr. **23**, 1249–1258 (1993).

Modal scintillation index: A physics-based statistic for acoustic source depth discrimination

V. Premus
MIT Lincoln Laboratory, 244 Wood Street, Lexington, Massachusetts 02420-9185

(Received 25 September 1997; revised 13 August 1998; accepted 10 December 1998)

A new discriminant based on the scintillation of normal mode amplitudes is introduced for the problem of passive surface/submerged source classification in a shallow water waveguide. The scintillation of modal energies is often used to characterize and understand acoustic wave propagation in a randomly fluctuating ocean waveguide [D. B. Creamer, J. Acoust. Soc. Am. **99**, 2825 (1996)]. This paper proposes a variant of the traditional modal scintillation index to treat the discrimination problem in a typical littoral oceanic waveguide. The approach is based on a modal decomposition of fluctuations in the received pressure field associated with the temporal modulation of the depth of an acoustic source about its mean value. Source depth fluctuations are the result of a platform's response to surface or internal wave motion. The rms mode excitations due to source depth modulation are shown to exhibit a depth dependent signature that may be exploited to statistically separate surface and submerged source classes. In this work, the modal scintillation index (SI) is defined as the variance in the estimated magnitude of the modal excitation normalized by its expected value over some observation interval. The statistic is self-normalizing, so knowledge of source level and source range is not required to separate the two source classes. Estimation of the modal excitation statistics requires only knowledge of the water depth and the sound speed profile at the array. Classification performance predictions in terms of receiver operating characteristic (ROC) curves will be presented based on KRAKEN Monte Carlo simulations under conditions of known and unknown source depth and range in spatially white Gaussian noise. An *ad hoc* decision criterion, which compares the minimum scintillation index across all modes to a threshold, was used to illustrate the phenomenology. Vertical line array and horizontal line array endfire geometries were considered. The modal scintillation approach may provide a robust alternative to matched field processing for the problem of binary source depth classification. © *1999 Acoustical Society of America.* [S0001-4966(99)02104-9]

PACS numbers: 43.30.Bp, 43.30.Wi [SAC-B]

INTRODUCTION

The scintillation of modal energies has often been used to characterize and understand acoustic wave propagation in a randomly fluctuating ocean waveguide.[1] In this paper, the modal scintillation index (SI), is introduced for the purpose of surface/submerged source discrimination in a shallow water waveguide.[2] The approach is based on the premise that temporal fluctuations in the depth of an acoustic source about its mean value are observable in the form of amplitude modulation of the mode excitations measured at the receiver array. Rms mode excitations associated with source depth modulation exhibit a depth dependent signature that may be exploited to statistically separate surface and submerged source classes. Source depth fluctuations are caused by a platform's response to surface or internal waves. Such depth fluctuations are generally of low frequency, much lower than that of the acoustic carrier. Like matched field processing (MFP), the modal scintillation classifier resolves depth dependent structure by integrating a physical model for acoustic propagation into the signal processing algorithm. However, the approach differs from MFP in that it directly casts the source classification problem as a binary hypothesis test.

The modal scintillation index is defined as the variance in the estimated modal excitation normalized by its expected value over some observation interval. The statistic is self-normalizing, so knowledge of source level and source range is not required to separate the two source classes. Estimation of the modal excitation statistics requires only knowledge of the water depth and the sound speed profile at the array.

The purpose of this paper is twofold. First, it will be demonstrated that mode amplitude fluctuations induced by source depth modulation are observable under nominal conditions of source level, noise spectrum level, and source/receiver range separation. Second, the statistical separation in the passive fluctuation signatures from surface and submerged sources for the same rms depth fluctuation will be quantified in terms of receiver operating characteristic (ROC) curves. An *ad hoc* decision criterion, which compares the minimum scintillation index across all modes to a threshold, is used to illustrate the phenomenology. The organization of this paper is as follows. In Sec. I, an overview of the method is presented, including the mathematical definition of the test statistic and a review of environmental model assumptions. In Sec. II, the results of a simulation experiment for a shallow water waveguide are presented. Classification ROC curves were derived from KRAKEN Monte Carlo simulations under conditions of known and unknown source depth and range in spatially white Gaussian noise. Results are presented for both a vertical line array and a horizontal line array at endfire. In Sec. III, we discuss the modal scintillation index based classifier in comparison with alternative techniques for source depth classification. In Sec. IV, results are summarized and future research directions are identified.

 0001-4966/99/105(4)/2170/11/$15.00 © 1999 Acoustical Society of America

I. THE MODAL SCINTILLATION INDEX

The adiabatic normal mode approximation has been used extensively by the matched field processing community to model low frequency, full-wave acoustic propagation in the ocean.[3] In the far field of an acoustic source, the complex pressure field at a depth z may be expressed in terms of a superposition of normal modes, $\Phi_m(z)$, given by[4]

$$p(z;z_s,r_s)=a\sum_{m=1}^{M}\Phi_m(z_s)\Phi_m(z)\frac{\exp(jk_{rm}r_s)}{\sqrt{k_{rm}r_s}}, \tag{1}$$

where z_s is source depth, r_s is source range, k_{rm} is the horizontal wave number corresponding to mode $\Phi_m(z)$, and a is a complex, zero-mean random variable with variance σ_a^2, corresponding to a narrow-band stochastic source signal. The source signal is assumed to have stable amplitude and phase over an interval, τ, such that

$$\tau \gg T, \tag{2}$$

where T is inversely related to the bandwidth of the source depth modulation process. In vector notation, temporal samples of the noise added pressure field, $\mathbf{p}(t)$, may be written as

$$\mathbf{p}(t)=\mathbf{\Phi}\mathbf{H}(t)+\mathbf{n}(t), \tag{3}$$

where $\mathbf{\Phi}$ is the $N\times M$ matrix of normal mode functions, $\mathbf{H}(t)$ is the $M\times 1$ vector of temporally varying modal excitations, $\alpha\Phi_m(z_s(t))[\exp(jk_{rm}r_s(t))/\sqrt{k_{rm}r_s(t)}]$, $\mathbf{n}(t)$ is an $N\times 1$ noise vector, and N is the number of sensors.

From Eq. (1), the pressure signal measured in the far field is directly proportional to the excitation level, $\Phi_m(z_s(t))$, of each mode at the depth of the source. Temporal estimates of the modal excitation time series are computed from samples of the received pressure field via the pseudoinverse calculation given by[5]

$$\hat{\mathbf{H}}(t)=\mathbf{\Phi}^{\dagger}\mathbf{p}(t), \tag{4}$$

where $\mathbf{\Phi}^{\dagger}=(\mathbf{\Phi}^H\mathbf{\Phi})^{-1}\mathbf{\Phi}^H$ represents the pseudoinverse of $\mathbf{\Phi}$, and the superscript H denotes conjugate transpose. Substituting Eq. (3) into Eq. (4) yields an estimate of the modal excitation time series,

$$\hat{\mathbf{H}}(t)=\mathbf{H}(t)+\mathbf{\Phi}^{\dagger}\mathbf{n}(t). \tag{5}$$

It should be pointed out that errors in the estimation of $\hat{\mathbf{H}}(t)$ associated with nonideal spatial sampling of the mode functions (e.g., nonfully spanning aperture, sensor position uncertainty, etc.) are not addressed in this work. In this work, it is assumed that the receiver array is capable of fully resolving the individual mode amplitudes. In a recent paper, Buck *et al.*[6] investigated Cramer–Rao bounds on the estimation of modal coefficients for several commonly used mode filtering techniques, including the pseudoinverse approach. They found that the pseudoinverse mode filter can be sensitive to errors in $\mathbf{\Phi}$ associated with numerical integration and environmental mismatch. This sensitivity is manifested in the form of a bias in the covariance of the mode coefficients, resulting from the fact that $\mathbf{\Phi}^{\dagger}\mathbf{\Phi}\neq\mathbf{I}$.[6] In the context of the modal scintillation classifier, this could lead to overestimation of the variance in mode amplitude fluctuations, and, in a worst case scenario, false dismissals of submerged sources as surface contacts. The pseudoinverse is one technique for the estimation of the modal coefficients from pressure field data. It is used here to facilitate the discussion of the methodology for depth discrimination based on the modal scintillation index. Other techniques for estimating $\hat{\mathbf{H}}(t)$ that may be more robust in the presence of imperfect environmental sampling may be substituted. In the future, it may be desirable to quantify the impact of estimation error on classifier performance predictions.

The scintillation index, SI_m, for each mode is computed from the estimated modal excitation time series in Eq. (5) via

$$\mathrm{SI}_m=\frac{\mathrm{Var}\{|\hat{h}_m(t)|\}}{E\{|\hat{h}_m(t)|\}}, \tag{6}$$

where $\hat{h}_m(t)$ represents the mth element of $\hat{\mathbf{H}}(t)$ and $|\cdot|$ denotes magnitude. The interval over which the statistics in Eq. (6) are calculated must be long enough to span a few cycles of the depth modulation process for estimate stability. The interval must also be short enough to insure that fine-scale ocean sound speed variability, as well as source amplitude and phase, are relatively stable over the averaging window. The modal scintillation index in Eq. (6) is based on the statistics of the mode excitation magnitudes. Thus it is expected to be insensitive to temporal variation in source range, r_s, subject to the condition that the spatial sampling at the receiver array preserves the orthogonality of the mode functions. Under this condition, the range dependent phases in the modal interference terms embedded in Eq. (4) do not contribute to the output of the mode filtering operation.

Figure 1 summarizes the modal scintillation concept. Figure 1(a) depicts the modal structure and sound speed profile typical of a downward refracting, shallow water environment in summer.[7] The channel supports three propagating modes at a source frequency of 35 Hz. In this paper, the focus is on a shallow water waveguide ($\sim$100–200 m) at low frequency ($<$50 Hz). Under these conditions, pressure field fluctuations due to random inhomogeneities of the medium are minimized, and sound propagation is well described with a small number of discrete modes. In this context, it becomes feasible to isolate the impact of the depth modulation process on the statistics of the individual mode excitations.

As stated earlier, for an acoustic source whose depth is fluctuating in response to wave interaction, the mode amplitude variance has a component which is a sensitive indicator of its mean depth. Mode amplitude fluctuations will exhibit high variance when the source is at or near the depth of a modal zero-crossing, where the derivative of the mode function is maximum [see Fig. 1(b)]. Similarly, mode amplitudes fluctuate with low variance for a source near a modal extremum, where the vertical derivative is zero [see Fig. 1(c)]. The critical property of the shallow waveguide which motivates the use of the modal scintillation index for binary depth classification is the fact that the normal modes are nearly sinusoidal and share a common zero-crossing at the pressure-

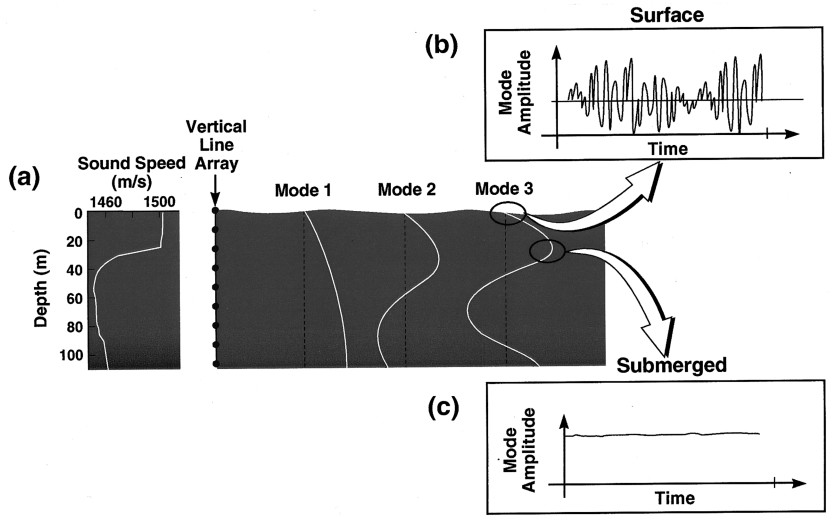


FIG. 1. Conceptual description of physics-based depth discrimination using the modal scintillation index: (a) Acoustic normal modes for downward refracting shallow water waveguide at 35 Hz. Modes exhibit sinusoidal depth dependence with common zero-crossing at pressure release surface. (b) Mode amplitude fluctuation associated with a surface source undergoing depth modulation in response to surface waves. Due to the large local derivative for all modes near the pressure release surface, source depth modulation near the surface maps to strong mode amplitude fluctuations for all modes simultaneously. (c) Mode amplitude fluctuation for submerged source undergoing similar magnitude depth modulation in response to internal waves. Due to the small local derivative of the mode function in the vicinity of an extremum, source depth modulation at this depth causes very little fluctuation in modal excitation.

release surface. This is the only part of the water column for which this property holds true. Thus a surface source with a given rms vertical motion will exhibit the high mode ''scintillation'' of Fig. 1(b) across **all** modes simultaneously. A submerged source with the same rms vertical motion will exhibit the very low mode scintillation of Fig. 1(c) for **at least one** mode, due to its expected proximity to at least one modal extremum.

Figure 2 shows estimated probability density functions (pdf)s of mode 3 scintillation index obtained from Monte Carlo simulation using 1000 trials under each hypothesis. The submerged source pdf is denoted $p(\mathrm{SI}|H_1)$. The surface source pdf is denoted $p(\mathrm{SI}|H_0)$. The source range in each case is 5 km. The pdf of the noise equivalent scintillation index is denoted by $p(\mathrm{SI}|n)$. Note from Eq. (5) that there will be an equivalent noise alone scintillation index when there is no acoustic source present. The modal scintillation index under the noise-only hypothesis, determined in a ''quiet'' neighboring frequency bin, will be helpful for calibration and *in situ* threshold setting. The pdf of the submerged source scintillation index always separates to the left of the background, in the direction of small values of SI. The pdf of the surface source scintillation index always separates to the right of the background, in the direction of large values of SI. If there is insufficient signal excess to observe the depth modulation-induced SI signal, both pdfs collapse back into the noise equivalent pdf.

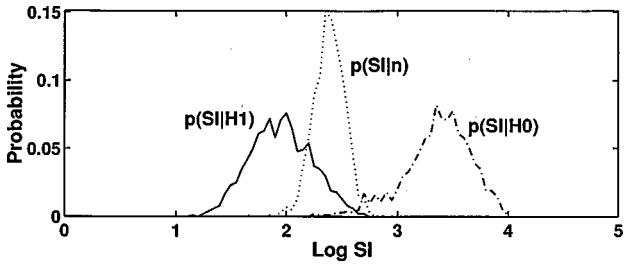


FIG. 2. Estimated probability density functions (pdf)s for mode 3 scintillation index derived from Monte Carlo simulation of 1000 trials under each hypothesis. The pdf of mode 3 SI for a submerged source (solid), $p(\mathrm{SI}|H_1)$, corresponds to a source at 25 m depth with a source level (SL) of 130 dB *re*: 1 μPa @ 1 m. The pdf of mode 3 SI for a surface source (dashed–dot), $p(\mathrm{SI}|H_0)$, corresponds to a source at 5 m depth with a SL of 160 dB *re*: 1 μPa @ 1 m. The pdf of the background scintillation index (dotted), $p(\mathrm{SI}|n)$, is associated with spatially white Gaussian noise at 65 dB *re*: 1 μPa in a 1-Hz band.

Throughout the paper, the depth modulation process under each hypothesis is modeled as a white Gaussian process with a prescribed standard deviation. It is important to emphasize that the model of the depth fluctuation process is always **exactly** the same for each source class. Discrimination based on the modal scintillation index exploits differences surrounding the way in which source depth fluctuations couple into the modal excitations. Depth discrimination using modal scintillation does not presuppose differences in the rms amplitude of the depth modulation process associated with each source class. Intuition would suggest that surface sources are likely to experience more significant depth excursions about their mean value than submerged sources due to interaction with surface waves. This attribute can only serve to enhance the scintillation index pdf separation between the two source classes. Submerged sources can also experience depth perturbations. Thus rms source depth fluctuation itself, if it were recoverable from the acoustic time series, is not expected to be a reliable statistic for discriminating between source classes.

II. SHALLOW WATER SIMULATION EXPERIMENT

To illustrate the utility of the modal scintillation index for source depth discrimination, a simulation experiment was performed for a shallow water waveguide. The simulation geometry is depicted in Fig. 3. Both vertical line array (VLA) and horizontal line array (HLA) at endfire geometries were considered. The depth of the channel is 110 m. The source frequency is 35 Hz. The sound speed profile for the experiment is depicted in Fig. 4. The solid curve denotes the environmental model used to compute the mode functions for the scintillation index calculation. The dash–dot curve denotes a single realization of the sound speed profile used to simulate the received pressure field. Simulated pressure field observations were generated from the perturbed sound speed profile to establish that the modal scintillation computation is relatively insensitive to small scale errors in the assumed environmental model. The bottom is modeled as a fluid half-

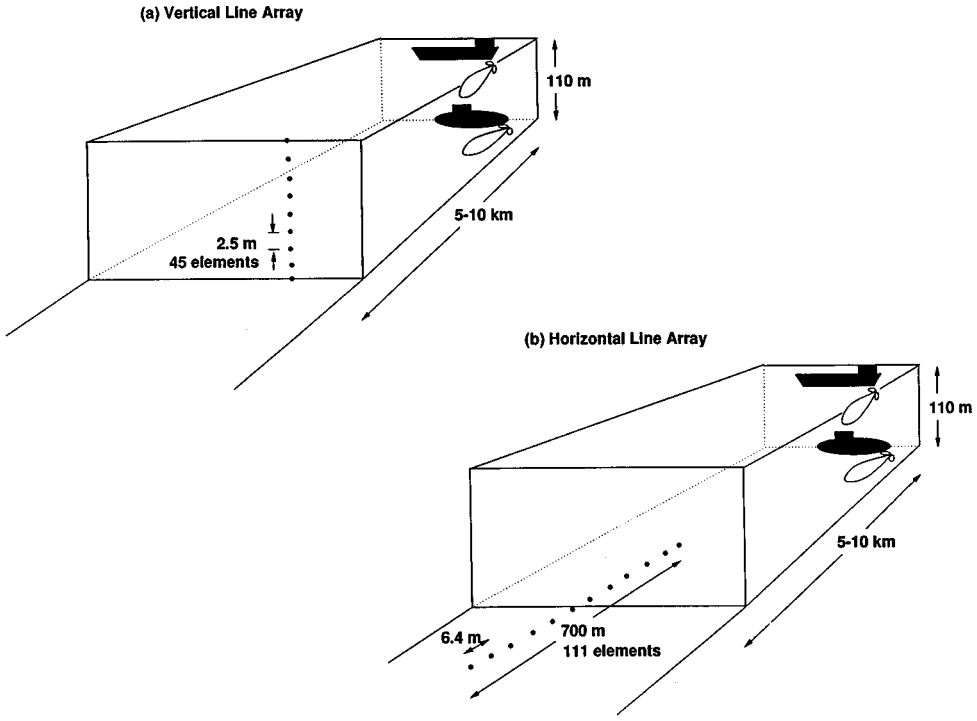


FIG. 3. Simulation geometry of modal scintillation experiment. Channel depth is 110 m. Source range is varied from 5 to 10 km during experiment. Surface source mean depth is 5 m. Submerged source mean depth is uniformly distributed from 20 m to 100 m. Standard deviation of source depth fluctuation is 1 m for each source class. (a) Fully spanning vertical line array with 45 sensors uniformly spaced at 2.5-m intervals. (b) Horizontal line array of length equal to 700 m, 45 sensors uniformly spaced at 16-m intervals. Target is located to endfire of horizontal array.

space with $c_b = 1800$ m/s and $\rho_b = 1.8$ g/cm^3. Throughout the simulation experiment, the noise field is modeled as spatially white Gaussian noise with a noise spectrum level, NL, of 65 dB *re*: 1 μPa in a 1-Hz band.

The VLA was a fully spanning array consisting of 45 hydrophones equally spaced at 2.5 m. The HLA consisted of 111 hydrophones uniformly spaced every 6.4 m in range over an aperture length, L_a, equal to 700 m. This choice of L_a represents the minimum array length which approximately satisfies the mode orthogonality condition over the aperture. This condition insures that the array is of sufficient length to resolve the two most closely spaced modes in wave number. For a horizontal aperture distributed in range, the mode orthogonality condition is given by

$$\int_0^{L_a} \Phi_m \Phi_n^* e^{j(k_{rm}-k_{rn})r}\, dr = 0, \quad m,n = 1,\ldots,M, \tag{7}$$

where r is range relative to the phase center of the array. It is straightforward to show that, for a rectangular shading function, this condition leads to a lower bound on L_a given by

$$L_a > \frac{2\pi}{\Delta k_{\min}}, \tag{8}$$

where $\Delta k_{\min}$ represents the difference between the two most closely spaced horizontal wave numbers supported by the channel. For this experiment, a KRAKEN calculation shows that $\Delta k_{\min}$ is approximately 0.007 m^{-1}. This yields an L_a of about 850 m. A plot of the respective modal innerproducts in

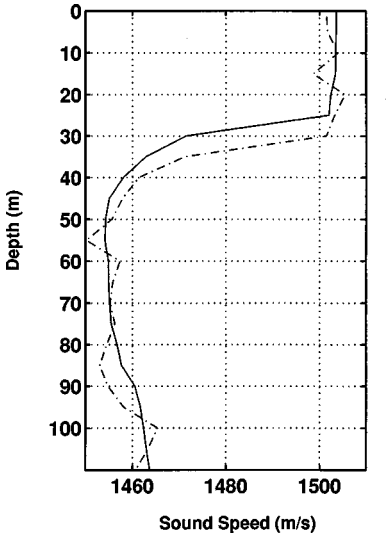


FIG. 4. Sound speed profile for simulation experiment. Solid curve denotes the environmental model used to compute mode functions for scintillation index calculation. Dash–dot curve denotes realization of sound speed profile used to simulate received pressure field.

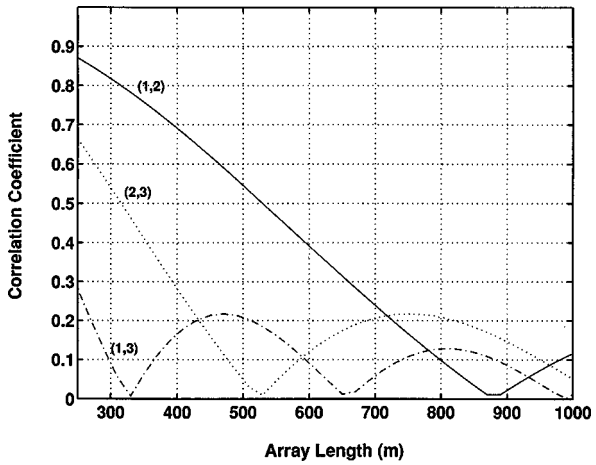


FIG. 5. Modal innerproducts versus array length, L_a, calculated from Eq. (8).

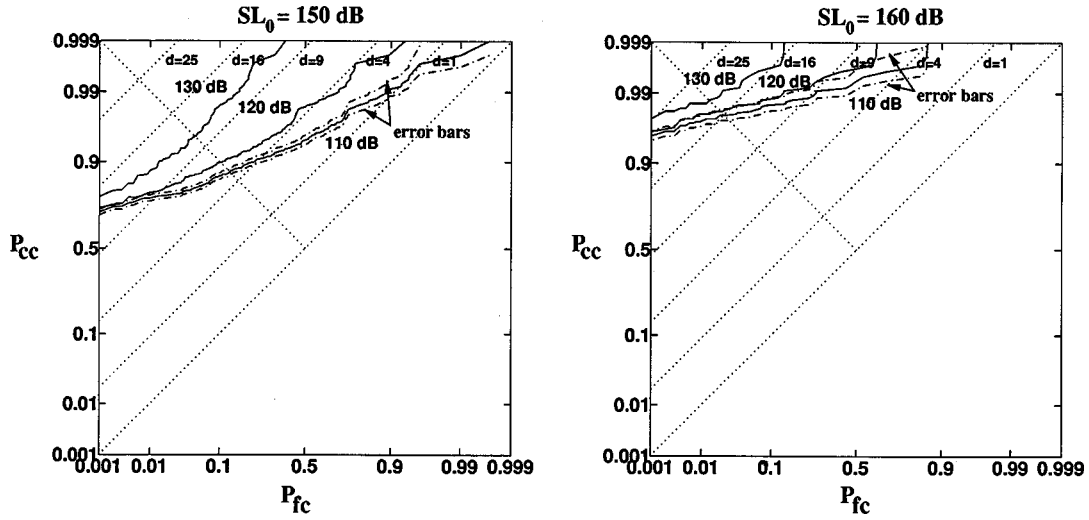


FIG. 6. Modal scintillation classification receiver operating characteristic (ROC) curves. Mean source depth assumed known under each hypothesis. Vertical line array. Illustrates performance of modal scintillation classifier as a function of increasing signal excess. Mode number is 3. Frequency is 35 Hz. Submerged source mean depth is 65 m. Surface source mean depth is 5 m. Rms source depth fluctuation is 1 m for each source class. Source range is 5 km for each class. SL_1 is varied from 110 dB to 130 dB. (a) SL_0 equals 150 dB. (b) SL_0 equals 160 dB. ROCs conditioned on the event that source is present with probability equal to 1. For a submerged source at 65 m depth, $SL_1=130$ dB corresponds to a signal excess of approximately 5 dB at the sensor, as measured in a 1-Hz band. For a surface source at 5 m depth, $SL_0=160$ dB corresponds to a signal excess of approximately 23 dB. The difference in signal excess at the receiver is accounted for by differences in transmission loss for the two source depths, about 12–13 dB.

Eq. (7) versus L_a, shown in Fig. 5, illustrates that for L_a as small as 700 m it should be possible to obtain modal resolution consistent with the condition of Eq. (8). The maximum array length is bounded above by limitations on spatial coherence of the received signal. Carey[8] has reported measurements for signal coherence in shallow water at 400 Hz on the order of 30λ. This value is likely to improve with decreasing acoustic frequency. At 35 Hz, L_a equal to 700 m corresponds to approximately 16λ, well within the bound reported by Carey.

Following Eq. (1), the source signal is modeled as a zero-mean, complex Gaussian random variable with a variance, σ_a^2, equal to $10^{SL/10}$, where SL denotes source level in units of dB *re*: 1 μPa @ 1 m. Temporal source depth fluctuation is modeled as a zero-mean, independent Gaussian random variable with a standard deviation of 1 m. As emphasized in the previous section, the rms amplitude of the depth modulation process is exactly the same under each hypothesis. For this experiment, a low-pass surface wave spectrum is assumed, with a high frequency cutoff of approximately 0.1 Hz. In order to satisfy the Nyquist criterion, and thereby avoid aliasing in the sampled surface wave fluctuation spectrum, the received pressure field is sampled every 5 s. The source signal in each case is assumed to have stable amplitude and phase over an interval on the order of 30–60 s. At 50% overlap, this interval is long enough to allow for a minimum of 20–30 independent samples of the pressure field fluctuation for each estimate of the modal scintillation index. For a surface source, denoted henceforth as H_0, the mean source depth is fixed at 5 m. For a submerged source, denoted henceforth as H_1, the mean source depth may range from 20 to 100 m.

A couple of comments pertaining to the ROC curve calculations that follow: ROC curves were computed via Monte Carlo simulation. Unless otherwise indicated, 1000 trials under each hypothesis were used. P_{cc}, the probability of correct classification, is defined as the probability of correctly deciding that the source is submerged. P_{fc}, the probability of false classification, is defined as the probability of incorrectly deciding the source is submerged when it is in fact surfaced. All ROC curves were computed under the assumption that a source is present with probability equal to 1. This is reasonable since, in practice, the classifier would be inserted into a sonar system as a post-processing step to an optimum detection algorithm. A brief, but more general, description of the calculation and interpretation of ROC curves in presented in the Appendix.

A. Source depth known exactly

First, we consider classifier performance for the case of mean source depth known exactly (DKE) under hypothesis H_1. The DKE condition implies that the mode which yields maximum pdf separation is known *a priori*. This condition enables an upper bound on classifier performance to be calculated as a function of source depth for source levels and noise levels of interest.

Figure 6 depicts classification ROC results for the VLA geometry. The mean H_1 source depth is 65 m. Mode 3 scintillation index is used to compute the ROC curves in this case since this source depth falls nearest to a maximum for that normal mode. The source range under each hypothesis is 5 km. Figure 6(a) and (b) correspond to H_0 source levels, SL_0, of 150 dB and 160 dB, respectively. In each case, classification performance is depicted for three values of H_1 source level, 110 dB, 120 dB, and 130 dB. Classification performance is good, particularly as the submerged source level approaches 130 dB. For the $SL_0=150$ dB case, extrapolating toward small P_{fc} along a trajectory parallel to the lines of the Gaussian ROC template (dotted lines), the classifier

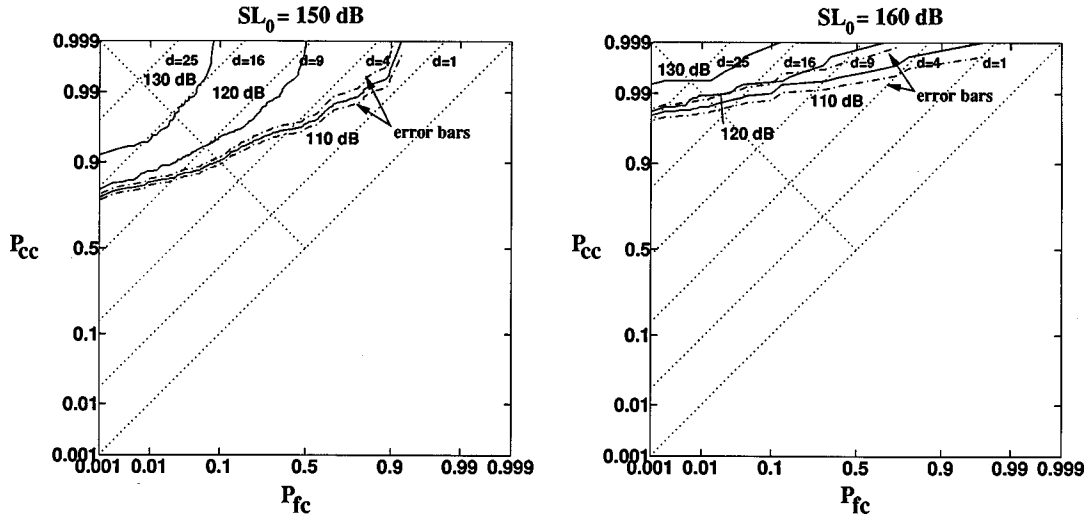


FIG. 7. Modal scintillation classification ROC curves. Mean source depth known under each hypothesis. Horizontal line array. Array length is 700 m, with 45 sensors spaced at 16-m intervals. Mode number is 3. Frequency is 35 Hz. Submerged source mean depth is 65 m. Surface source mean depth is 5 m. Rms source depth fluctuation is 1 m for each source class. SL_1 varied from 110 dB to 130 dB. (a) SL_0 equals 150 dB. (b) SL_0 equals 160 dB. ROCs conditioned on the event that source is present with probability equal to 1.

performs with a P_{cc} in excess of 0.5 for a false classification probability of 0.001. When SL_0 is increased by 10 dB, the classifier operating point improves to a P_{cc} of about 0.95 for the same P_{fc}. Thus as one would expect, classification performance improves significantly with signal excess under either hypothesis. Signal excess can be estimated from a power spectral density estimate at the receiver. For reference, the submerged source with a SL of 130 dB at a range of 5 km in this shallow water environment corresponds to a signal excess of about 5 dB at the sensor (no coherent processing gain due to hydrophone grouping is assumed). A surface source with a SL of 150 dB at a range of 5 km corresponds to a signal excess of about 13 dB at the sensor. The 12-dB difference in signal excess between the two classes is accounted for by the additional transmission loss incurred by a source near the pressure release surface.

Figure 7 depicts the corresponding classification ROC curve results for the HLA endfire geometry under the same simulation conditions as in Fig. 6. At this source depth, the capacity of the HLA to separate the two source classes based

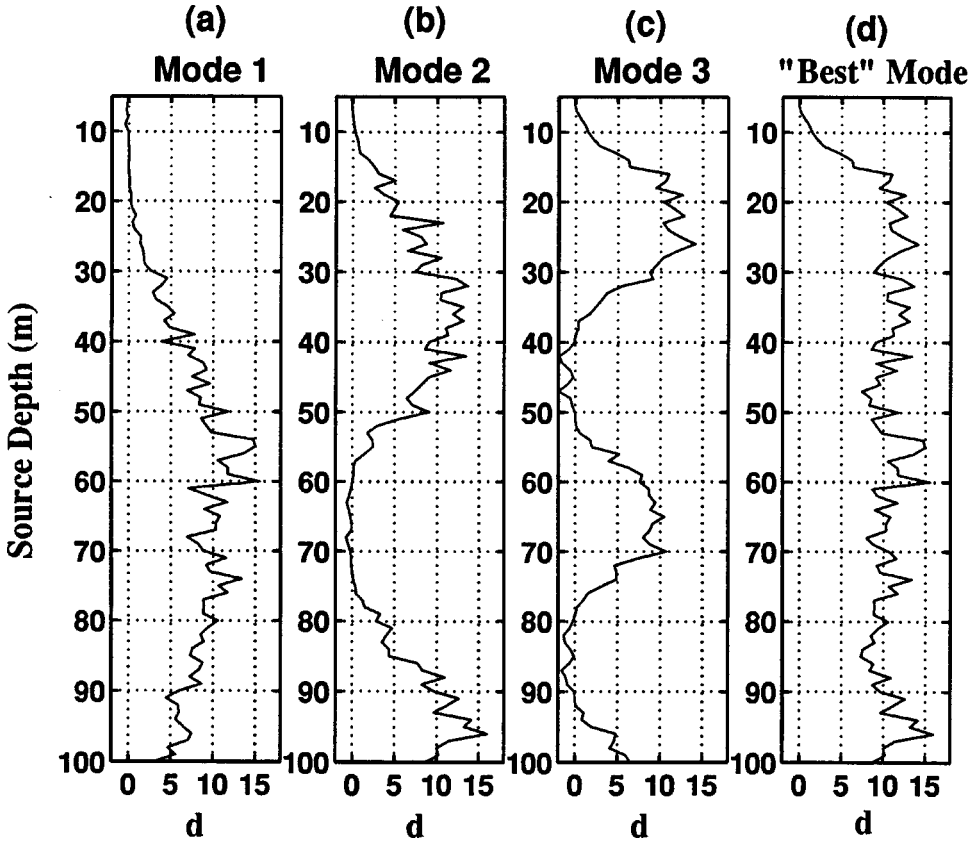


FIG. 8. Source depth dependence of scintillation index pdf separation. Vertical line array. Frequency is 35 Hz. SL is 130 dB *re*: 1 μPa @ 1 m. NL is 65 dB *re*: 1 μPa in a 1-Hz band. Depth modulation is modeled as an independent Gaussian random variable with zero mean and standard deviation equal to 1 m. Pdf separation quantified in terms of deflection index, *d*, as defined in Eq. (10). Uniform $H1$ classification coverage is obtained for source depths exceeding the shallowest modal maximum. $H1$ modal scintillation appears most ''clutter-like'' near modal zero-crossings. Number of Monte Carlo trials at each depth is 100.

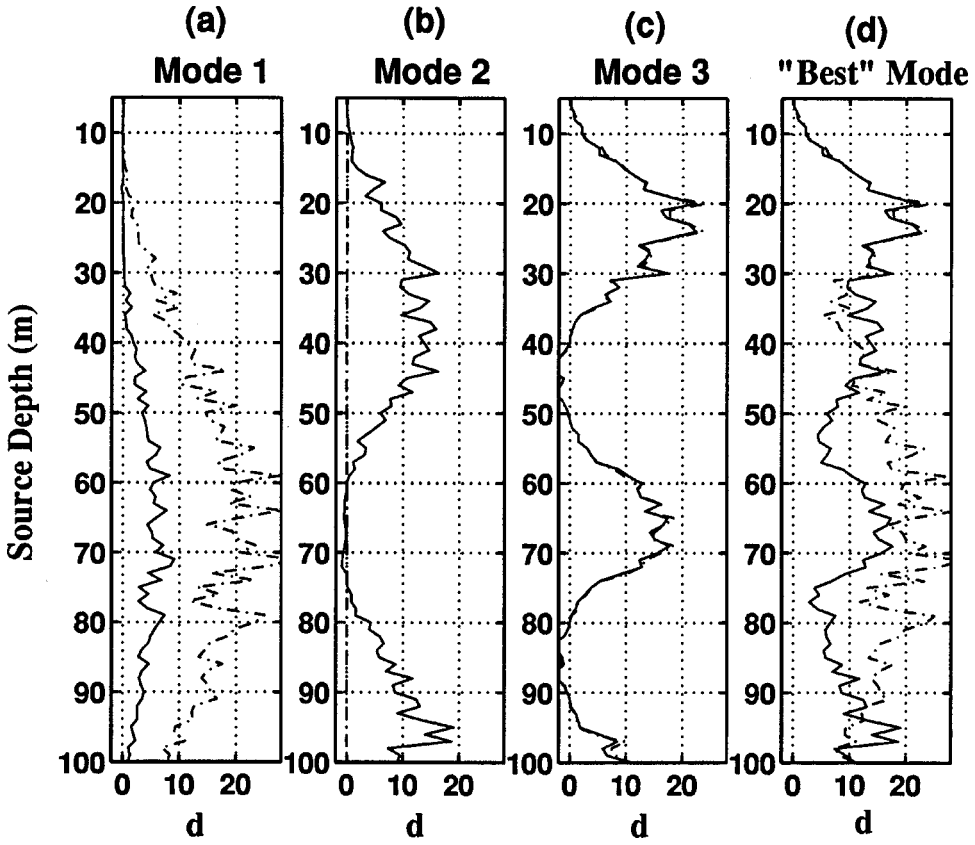


FIG. 9. Source depth dependence of scintillation index pdf separation. Horizontal line array. Array length is 700 m, with 45 sensors spaced at 16-m intervals. Frequency is 35 Hz. SL is 130 dB *re*: 1 μPa @ 1 m. NL is 65 dB *re*: 1 μPa in a 1-Hz band. Depth modulation is modeled as an independent Gaussian random variable with zero mean and standard deviation equal to 1 m. Pdf separation quantified in terms of deflection index, d, as defined in Eq. (10). Horizontal line array classification performance is sensitive to depth of the array. Solid curves correspond to $d(z_s)$ for a bottom-mounted array at a depth of 110 m. Dashed curves correspond to $d(z_s)$ for an array depth of 67 m. Number of Monte Carlo trials at each depth is 100.

on mode 3 scintillation index is much like that of the fully spanning VLA. The slight improvement in the HLA ROC performance over that of VLA geometry is attributable to the fact that three times as many sensors were employed in the HLA configuration.

B. Source depth dependence of pdf separation

Figure 8 illustrates the source depth dependence of the modal scintillation index pdf separation for the VLA geometry. The purpose of this part of the simulation experiment was to identify the depth at which the two source classes begin to separate, and to quantify the expected variation of classifier performance throughout the entire water column. The metric used to quantify pdf separation was the detection index, d, defined by[9]

$$d=\frac{(E\{\mathrm{SI}|H_1\}-E\{\mathrm{SI}|H_0\})^2}{\sigma^2_{\mathrm{SI}|H_0}}. \tag{9}$$

d was computed for each mode based on 100 trials at each source depth between 5 m and 100 m in increments of 1 m. The source level was held fixed at 130 dB for all source depths. The H_0 data were sampled at a source depth equal to 5 m (hence $d=0$ at that source depth for all modes). Figure 8(a)–(c) depict $d(z_s)$ for each of the three modes supported by the waveguide at the source frequency of 35 Hz. The results show that the mode which presents the best classification opportunity is a strong function of mean source depth. An *ad hoc* decision rule which compares the minimum scintillation index across all modes to a threshold, shown in Fig. 8(d), yields uniform classification coverage versus depth for source depths exceeding approximately 16 m. This depth constitutes the boundary between the surface and submerged source classes based on the minimum modal scintillation index criterion. It is expected to occur at a depth slightly less than the shallowest modal maximum supported by the waveguide at the source frequency.

When the mean source depth approaches a depth corresponding to a zero-crossing (for example, Mode 3 at 45 m) the scintillation index for that mode appears most ''clutterlike.'' Clutterlike in this context is defined by the condition $d(z_s)\sim 0$. A depth-modulated source in this part of the water column presents a potential false classification opportunity due to the poor separation between the H_0 and H_1 pdfs. However, if the source frequency is high enough to support the propagation of multiple normal modes, it is expected that the distribution of modal maxima in depth will be sufficient to insure that the *ad hoc* decision rule yields an acceptable classification performance level.

Figure 9 depicts the source depth dependence of the modal scintillation index pdf separation for the HLA endfire geometry. Detection index is plotted for two array depths, 67 m (dash–dot) and 100 m (solid). Since the pressure field at the receiver is proportional to the modal excitation level at the depth of the sensor, classification performance for the HLA is expected to be strongly dependent on array depth. For the bottom mounted array at $z=110$ m, mode 1 is excited at only 50% of its maximum amplitude at the depth of the array. Thus all source depths in Fig. 8 for which the mode 1 scintillation index determined the classification performance under the VLA geometry incur a degradation in surface/submerged pdf separation under the HLA geometry. Inspection of the composite performance given in Fig. 9(d) illustrates that the affected source depths lie in the ranges

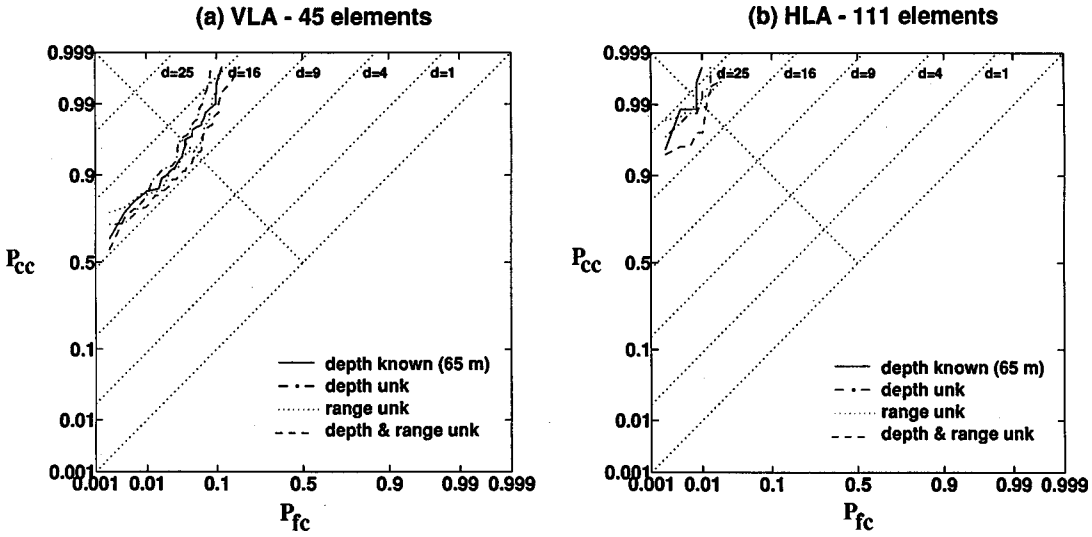


FIG. 10. Modal scintillation classification ROC curves for the case of unknown source depth and range. *Ad hoc* decision rule compares minimum scintillation index to a threshold. (a) Vertical line array, (b) Horizontal line array. Conditions common to all curves are: $SL_1=130$ dB, $SL_0=160$ dB, NL=65 dB *re*: 1 μPa in a 1-Hz band, $z_s=5$ m under H_0, and $\sigma_z=1$ m under each source class. For both (a) and (b): (Solid) Known source depth under H_1 corresponding to $z_s=65$ m and $r_s=5$ km. (Dash–dot) Unknown source depth under H_1 with mean submerged source depth uniformly distributed from 20 m to 100 m. (Dotted) Unknown range under each hypothesis, and known depth under H_1 with source range under each hypothesis sampled from two independent uniform distributions between 5 km and 10 km. (Dashed) Unknown source depth under H_1 and unknown source range under each hypothesis with sampling conditions in each case as described above. ROCs conditioned on the event that source is present with probability equal to 1. In each case, ROC curves derived from Monte Carlo simulation of 1000 trials.

50–60 m and 75–85 m. For the mid-water column towed array at $z=67$ m, the array is positioned precisely at the mode 2 zero-crossing depth. Thus the source depth-modulation signature is totally undetectable with the mode 2 scintillation index. All source depths for which mode 2 scintillation index determined the classification performance incur a performance degradation relative to that realized under the VLA geometry. In this case, however, the range of affected source depths is limited to 30–40 m. Therefore, it is possible to mitigate this type of performance degradation with careful selection of the array depth, or control of the array depth in the case of a towed array.

C. Unknown source depth and range

The source depth dependence results of the previous section suggest that the minimum scintillation index criterion will yield effective classification performance when the source depth under H_1 is not known *a priori*. In this section, we examine ROC performance when the data are sampled from populations of unknown source depth under H_1 and unknown range under each hypothesis. Again, the decision rule which compares the minimum scintillation index across all modes to a threshold is used.

Figure 10(a) and (b) show the ROC results for both the vertical line array and horizontal line array at endfire, respectively. Simulation conditions common to all ROC curves are: $SL_1=130$ dB, $SL_0=160$ dB, NL=65 dB *re*: 1 μPa in a 1-Hz band, $z_s=5$ m under H_0, and $\sigma_z=1$ m under each source class. ROC curves were derived from Monte Carlo simulation of 1000 trials. For each array geometry, an example of known source depth under H_1 (solid) is reproduced from Fig. 8. In this case, $z_s=65$ m and $r_s=5$ km. This prior information directs us to use the mode 3 scintillation index to compute the ROC curve. Next, we consider the condition of unknown source depth under H_1 (dash–dot). In this case, the mean submerged source depth was sampled from a uniform distribution between 20 m and 100 m. The case of unknown range under each hypothesis and known depth under H_1 follows (dotted). This condition is simulated by sampling source range under each hypothesis from two independent uniform distributions between 5 km and 10 km. Last, the two uncertainty conditions, unknown source depth under H_1 and unknown source range under each hypothesis, are combined (dashed). A comparison of the four curves under each geometry shows that performance of the classifier is largely unaffected by the need to search mode space for the proper scintillation index when the source depth is not known *a priori*. The resulting ROC curves represent, in effect, the average performance based on a Monte Carlo integration over the uncertain variable(s) of interest. In the case of the VLA, it appears that the integration over uncertain depth (dash–dot) actually yielded slightly better performance than the depth known exactly case (solid) for a source depth of 65 m. Integration over uncertain depth and range causes a small dropoff in classification performance, due to the increased transmission loss associated with those Monte Carlo samples corresponding to ranges in excess of 5 km. From this numerical result, we draw the conclusion that the performance of the modal scintillation based classifier is not sensitive to uncertainty regarding source range and depth.

III. DISCUSSION

A. Background mode scintillation due to surface waves

Application of the mode scintillation classifier to data will require the identification of potential sources of back-

ground scintillation, not associated with source depth modulation, which could produce false classification results. One anticipated source of background mode scintillation is waveguide distortion due to surface waves. The fact that surface waves could elicit pressure field fluctuations in a shallow waveguide was first observed by Scrimger.[10] That author measured acoustic signal fluctuations between a fixed cw source and a fixed receiver by causing a bow wave from a small boat to pass over the receiver, transmission path, and source, respectively. The source frequency was in the range 1–10 kHz. The channel depth was approximately 3 m. Thus the relationship between acoustic wavelength and waveguide depth for Scrimger's experiment was similar to that adopted in this paper. Simultaneous measurements of the received signal and surface wave height above the source and receiver were recorded. Scrimger found that fluctuations in signal level occurred when the surface wave packet was above the source or receiver, and fluctuations were small when the surface wave packet was entirely between the source and receiver. These fluctuations are the result of the expansion and compression of the waveguide normal modes resulting from changes in the effective depth of the waveguide due to the surface wave disturbance. Such waveguide depth modulation in the vicinity of the receiver array could introduce a bias in the modal scintillation signature of a submerged source, potentially leading to a misclassification result. In practice, it would be necessary to subtract the mode scintillation component associated with surface waves over the receiver array. This could be accomplished by recording hydrostatic pressure, or equivalently, instantaneous water depth, at the receiver array with each pressure field snapshot. Scrimger's observation also suggests that it may be possible to classify a source fixed in depth by exploiting waveguide distortion induced by surface waves propagating over the source.

B. Alternative techniques for depth classification

The notion of exploiting statistical fluctuations in the received pressure field for the purpose of acoustic source classification is not new. Wagstaff[11] proposed an energy detection approach based on a sliding window harmonic mean calculation. The harmonic mean, $[\Sigma_{j=1}^{J} X_{jk}^{-1}]^{-1}$, where j is the time index and X_k represents the power spectrum estimate in frequency bin k, has been shown to enhance the detectability of stable amplitude signals relative to a fluctuating background. However, its use as a classifier is limited by the fact that amplitude stability is an attribute that is not necessarily unique to a particular source class. Simanin[12] appears to have been the first to suggest a propagation physics-based approach to the binary classification problem. Simanin proposed the discrimination of acoustic rays as ''water or surface rays'' based on statistical models for pressure amplitude fluctuations due to propagation through water column inhomogeneities.[12] The modal scintillation approach is distinguished from these earlier fluctuation-based classification techniques in that it incorporates a model for the propagation physics directly into the discrimination algorithm. Consequently, it is able to operate on information related to the depth dependent coupling of dynamic variations in source depth to the waveguide without invoking any special assumptions (e.g., phase stability) concerning the source signal under one hypothesis.

The modal scintillation classifier attempts to resolve depth dependent structure in the received field by integrating a physical model for waveguide propagation into the discrimination algorithm. In this respect, the technique can be viewed as a variant of matched field processing (MFP). Recent advances in MFP have also begun to consider the possibility of exploiting temporal source dynamics for improving source localization. For example, Tantum and Nolte[13] have demonstrated the use of matched field tracking which employs a Markov model for source motion based upon state transition probabilities. For the purpose of binary source depth classification, however, the modal scintillation approach may offer some advantages over a fully coherent MFP approach. First, it simplifies the objective by casting the source localization problem as a binary hypothesis test. Second, it redefines the signal of interest. By performing a modal decomposition and integrating the energy in temporal mode amplitude fluctuations, it addresses the classification problem within an incoherent framework. By discarding phase information, the modal scintillation approach may represent a robust alternative to coherent MFP for cases where the depth-modulation signal is observable. Last, by comparing the energy in modal excitations to a threshold, rather than computing a range/depth posterior probability map, classification is achievable without the computational burden associated with the exhaustive computation of replica field realizations common to many MFP implementations. It should be pointed out, however, that the modal scintillation classifier has more stringent spatial sampling requirements than MFP, due to the fact that it must resolve individual mode excitations.

IV. CONCLUSION

In this paper, the modal scintillation index has been introduced as a discriminant for source depth classification in a shallow water waveguide. The approach was motivated by the oceanic waveguide's sensitivity to temporal source depth dynamics associated with a platform's response to surface or internal wave motion. The modal scintillation classifier exploits differences in the ways in which source depth modulation couples to the waveguide for surface and nonsurface acoustic sources. This coupling is controlled by the local vertical derivative of the waveguide normal mode about the mean depth of the source. A source near the pressure release surface of a downward refracting shallow waveguide possesses the unique attribute that for a given rms vertical motion, all mode excitations will simultaneously exhibit a high valued scintillation index. This attribute persists for source depths down to the shallowest modal extremum supported by the waveguide at the source frequency, approximately 20 m for a 100-m waveguide at 35 Hz. This result follows from the sinusoidal nature of the normal modes in this environment, and the fact the pressure release surface is the only part of the water column for which all of the normal modes share a common zero crossing.

The modal scintillation approach differs from past attempts at fluctuation-based classification by virtue of its incorporation of a physical model for waveguide propagation. In this respect, the modal scintillation classifier can be viewed as a form of incoherent matched field processing. However, the approach differs from MFP by directly casting the classification problem as a binary hypothesis test. It thus avoids the potentially prohibitive computation of replica fields which are only of use when there is a depth resolution requirement. Further, due to its incoherent nature, emphasizing the relative energies in the peaks of the horizontal wave number spectrum, it may offer a more robust alternative to fully coherent MFP.

Simulation results have been presented in terms of classification ROC curves for both a vertical line array and a horizontal line array at endfire. Conditions of known and unknown source depth, and known and unknown source range in spatially white Gaussian noise were considered. An *ad hoc* decision criterion was employed, which compared the minimum modal scintillation index across all modes to a threshold. However, with prior knowledge of the mode functions, it is possible to analytically derive a likelihood ratio test for depth discrimination. Future work will be directed toward the experimental validation of source depth discrimination based on modal scintillation. It is expected that sensitivity of the approach to a number of important limiting environmental factors will have to addressed. These will include the impact of a nonfully spanning vertical aperture, uncertain sensor position, and the effects of mode coupling from range dependent bathymetry.

ACKNOWLEDGMENTS

The author gratefully acknowledges helpful technical discussions with Tom Green, Paul Kolodzy, Richard Lacoss, Dan Dudgeon, and Bill Payne of MIT Lincoln Laboratory, and Jim Lynch of Woods Hole Oceanographic Institution. This work was sponsored in part by SPAWAR, under Air Force Contract No. F19628-95-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the U. S. Government.

APPENDIX: RECEIVER OPERATING CHARACTERISTIC CURVES: CALCULATION AND INTERPRETATION

The use of the receiver operating characteristic (ROC) curve as a decision theoretic performance metric was first proposed by Peterson *et al.*[14] The ROC curve relates a decision algorithm's performance to the separation between the output probability density functions associated with two or more competing subject hypotheses. In the context of this paper, the term decision algorithm refers to a classifier. The classification ROC curve is then a parametric plot of correct classification probability, P_{cc}, versus false classification probability, P_{fc}, with the free parameter being the decision threshold of the classifier. In the statistics literature, the probability of correct classification is often referred to as the ''power'' of the test, while the false classification probability

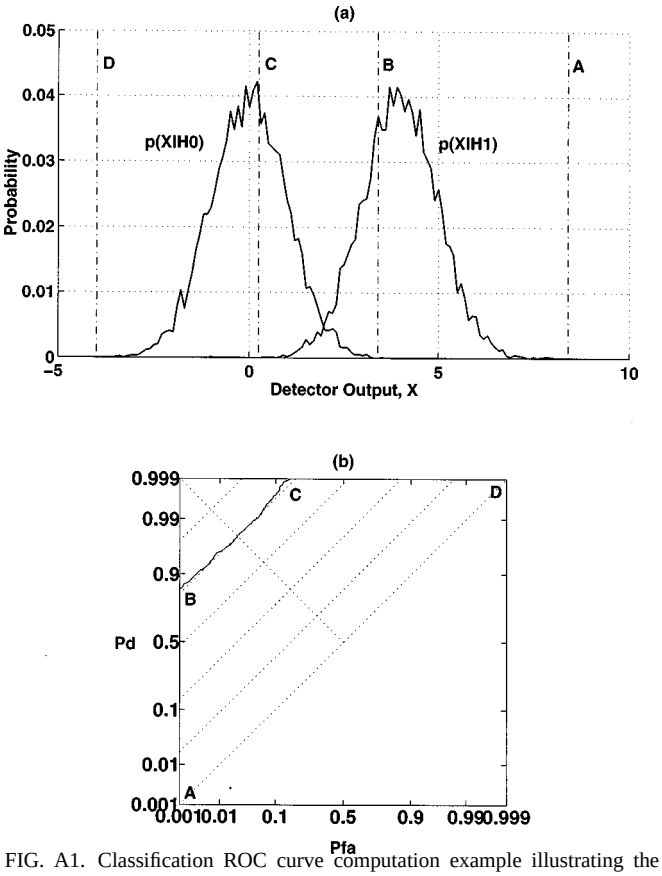


FIG. A1. Classification ROC curve computation example illustrating the mapping from probability space to (P_{cc}, P_{fc}) space: (a) probability densities of classifier output under each hypothesis, (b) classification ROC curve. As the threshold in (a) is decreased from A to D, the operating point on the ROC curve traverses from ''southwest'' to ''northeast.''

is referred to as the ''size'' of the test. It is clearly desirable to have powerful tests of small size.[15] Mathematically stated, the coordinates of any point on the ROC curve are obtained from the underlying probability densities of the two hypotheses by

$$P_{cc} = \int_{\gamma}^{+\infty} p_{X|H_1}(x|H_1)\,dx, \tag{A1}$$

and

$$P_{fc} = \int_{\gamma}^{+\infty} p_{X|H_0}(x|H_0)\,dx, \tag{A2}$$

where $p_{X|H_1}(x|H_1)$ is the pdf of the classifier output under hypothesis H_1, $p_{X|H_0}(x|H_0)$ is the pdf of the classifier output under hypothesis H_0, and γ is the threshold. Thus in order to trace out the entire ROC curve for a particular classifier, the threshold γ must be varied over the expected range of possible classifier output values.

To elucidate the interpretation of the ROC curve, it is useful to examine the mapping from pdf coordinates to (P_{cc}, P_{fc}) coordinates for a hypothetical experiment. Figure A1(a) depicts the probability densities underlying the two competing hypotheses in this experiment. The points labeled A–D in Fig. A1(a) represent different threshold settings. Figure A1(b) shows the resulting ROC curve for the classifier in this experiment. The (P_{cc}, P_{fc}) operating points corresponding to the four threshold settings are labeled. First, consider

threshold setting A. Observe from Fig. A1(a) that the probability of the classifier output exceeding threshold A is zero under either hypothesis. Consequently, this point maps to the origin of the (P_{cc}, P_{fc}) coordinate system. As the classifier threshold is lowered from A to B, probability density begins to accumulate under the H_1 hypothesis, while the probability of classifier output exceeding threshold B under the H_0 hypothesis remains zero. Thus the ROC curve departs the origin along the positive P_{cc} axis until it reaches B. As the decision threshold is further lowered from B to C, probability density is accumulated under each hypothesis. The trajectory of the receiver operating curve subsequently moves off of the P_{cc} axis. At C, all of the probability density under hypothesis H_1 has been accumulated, and P_{cc} attains its maximum value of 1. Finally, as the threshold is further reduced from C to D, all of the probability density under each hypothesis has been accumulated, corresponding to the point $P_{cc}=1$, $P_{fc}=1$ on the ROC coordinate axes. In short, classifier performance improves as the ROC curve tends toward the ''northwest'' corner of the (P_{cc}, P_{fc}) coordinate frame, where correct classification probability is maximized for a given value of false classification probability.

[1] D. B. Creamer, ''Scintillating shallow-water waveguides,'' J. Acoust. Soc. Am. **99**, 2825–2838 (1996).

[2] V. Premus, ''A mode space energy detector for acoustic source depth discrimination,'' J. Acoust. Soc. Am. **101**, 3026–3027 (1997).

[3] J. Krolik, ''Matched-field minimum variance beamforming in a random ocean channel,'' J. Acoust. Soc. Am. **92**, 1408–1419 (1992).

[4] F. Jensen, W. Kuperman, M. Porter, and H. Schmidt, *Computational Ocean Acoustics* (AIP Press, Woodbury, NY, 1994).

[5] G. R. Wilson, R. A. Koch, and P. J. Vidmar, ''Matched mode localization,'' J. Acoust. Soc. Am. **84**, 310–320 (1988).

[6] J. R. Buck, J. C. Preisig, and K. E. Wage, ''A unified framework for mode filtering and the maximum *a posteriori* mode filter,'' J. Acoust. Soc. Am. **103**, 1813–1824 (1998).

[7] W. M. Carey, J. Doutt, R. B. Evans, and L. M. Dillman, ''Shallow-water sound transmission measurements on the New Jersey continental shelf,'' IEEE J. Ocean Eng. **20**, 321–336 (1995).

[8] W. M. Carey, ''Recent shallow water measurements of signal coherence and array performance in several shallow water areas,'' J. Acoust. Soc. Am. **101**, 3030 (1997).

[9] A. D. Whalen, *Detection of Signals in Noise* (Academic, New York, 1971).

[10] J. A. Scrimger, ''Signal amplitude and phase fluctuations induced by surface waves in ducted sound propagation,'' J. Acoust. Soc. Am. **33**, 239–247 (1961).

[11] R. A. Wagstaff, ''Exploitation of fluctuations to enhance target detection and to reduce clutter and background noise in the marine environment, Proc. SPIE Aerospace/Defense Sensing and Controls '96 Symp. Targets and Backgrounds: Characterization and Representation II, **2742**, 276–283 (1996).

[12] A. A. Simanin, ''Possibility of classifying water and surface rays by amplitude fluctuations of the received signal,'' Sov. Phys. Acoust. **36**, 622–624 (1990).

[13] S. L. Tantum and L. W. Nolte, ''Tracking and localizing a moving source in an uncertain shallow water environment,'' J. Acoust. Soc. Am. (submitted).

[14] W. W. Peterson, T. G. Birdsall, and W. C. Fox, ''The theory of signal detectability,'' Trans. I.R.E. Professional Group Inf. Theory, **PGIT-4**, 171–212 (1954).

[15] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis* (Addison-Wesley, 1991).

Effects of salt on bubble acoustic radiation in water

Ali R. Kolaini
Mechanical and Environmental Engineering Department, University of California, Santa Barbara, Santa Barbara, California 93106

(Received 26 March 1998; revised 30 November 1998; accepted 14 December 1998)

Recent measurements of laboratory-generated noise by breaking waves exhibit an increase in sound-pressure levels for salt water compared to fresh water over a broader range of frequencies [A. R. Kolaini, J. Acoust. Soc. Am. **103**, 300–308 (1998)]. The surprise increase in sound-pressure levels inspired the study of the effects of salt on the sound radiation by single bubbles released from various size needles. A needle assembly and a false tank (2×2×4 ft), with acoustically transparent walls placed in the middle of an anechoic tank (12×12×8 ft), were used to study the acoustical characteristics of bubbles released from needles. The false tank was filled with water that contained various percentages of sodium chloride (NaCl). In this paper, the results of experiments to examine the variations in radiated sound pressure and the change in the damping coefficient of bubbles as the salinity level increased are reported. The effects of a surfactant agent such as Triton™ 100-X and other mineral acids such as HCl on the bubble acoustic radiation are discussed. The observations show that the change in the local surface tension may not alter the acoustic radiation, whereas the local influence of the ions on water structure, possibly related to the electric repulsion, may play a dominant role in altering the sound pressure and significantly reducing the quality factor of the bubble sound. © *1999 Acoustical Society of America.* [S0001-4966(99)06003-8]

PACS numbers: 43.30.Nb, 43.30.Es [SAC-B]

INTRODUCTION

It has been recognized that breaking waves and rainfall are the main source of ambient noise in the ocean over certain frequencies. Recently, much attention has been devoted to the understanding of the source mechanisms for ambient noise production at the sea surface, and has lent further support to the contributions of gas bubbles as the principal source of this noise.[1] However, a detailed description of the specific roles of these bubbles has not yet been given.[2–4] Investigation of the acoustic emissions by bubbles released from needles was studied by several authors.[5] The emission is that of a damped sinusoidal oscillator. The small-amplitude oscillation is modeled by using the Rayleigh–Plesset equation. The bubble dynamic equation has so far excluded the effects of the impurity in the host fluid.

The absence of bubble coalescence in salt water, which leads to significant change in bubble distribution, has been reported by numerous researchers, for example see Scott[6] and Monahan and Zeitlow.[7] The process of the bubble non-coalescence was ascribed by Kitchner[8] to be related to the existence of organic active materials that adsorbed on water–air interfaces, which would reduce the surface tension and in turn prevent thin films from rupturing. Another mechanism was suggested by Pounder[9] to be the preferential ion attachment on the bubble surface due to sodium chloride present in sea water that contributes an electric repulsion which tends to keep bubbles from coalescing. The effect of salt in reducing bubble coalescence in aqueous solutions has been noted in many oceanic and froth flotation systems.[10] The reasons for the effect of salt in reducing bubble coalescence and enhancing the bubble stability are not yet known. Recently, it has been shown that some salts and mineral acids have no effect on this phenomenon. Craig *et al.*[11] have shown that the electrolytes inhibiting coalescence have a correlation with the ionic strength, without offering an obvious explanation.

Recently, the laboratory bubble cloud generated by a tipping bucket by Carey *et al.*[12] and bubble-plume generation from a transient cylindrical water jet by Kolaini *et al.*[13] have shown a 3–4-dB reduction in sound-pressure level from the cloud in salt water compared to fresh water. They attributed the level reduction to enhanced thermal damping, coupled with an increase in the air/water interface area as was hypothesized by Prosperetti.[14] Despite recent studies of the effect of salt on bubble distribution, an accurate understanding of its effects on the radiated sound pressure from bubbles has not yet been given. The prime concern in this study is the effect of salts in the bubble dynamics and subsequent acoustic radiation. In this paper, we report the preliminary results of the bubble dynamics and subsequent acoustic radiation that are affected by salts. Even though these observations are preliminary at this stage, they may provide some valuable information about the influence of salts and organic compounds on the nature of acoustic radiation by bubbles, which may have direct consequences in the understanding of the ambient-noise generation in the ocean.

I. EXPERIMENTAL SETUP AND PROCEDURES

The apparatus used in this experiment consists of a bubble maker/hydrophone assembly with various size hypodermic needles, a B&K 8103 hydrophone, a 2×2×4 ft false tank with acoustically transparent walls submerged in an anechoic larger tank with dimensions 12×12×8 ft (see Fig. 1). The hydrophone was connected to a B&K charge amplifier and was bandpass filtered in the frequency range of 500–5000 Hz. The recorded signatures were then downloaded to a Macintosh computer via a LABVIEW program for further

 0001-4966/99/105(4)/2181/6/$15.00 © 1999 Acoustical Society of America

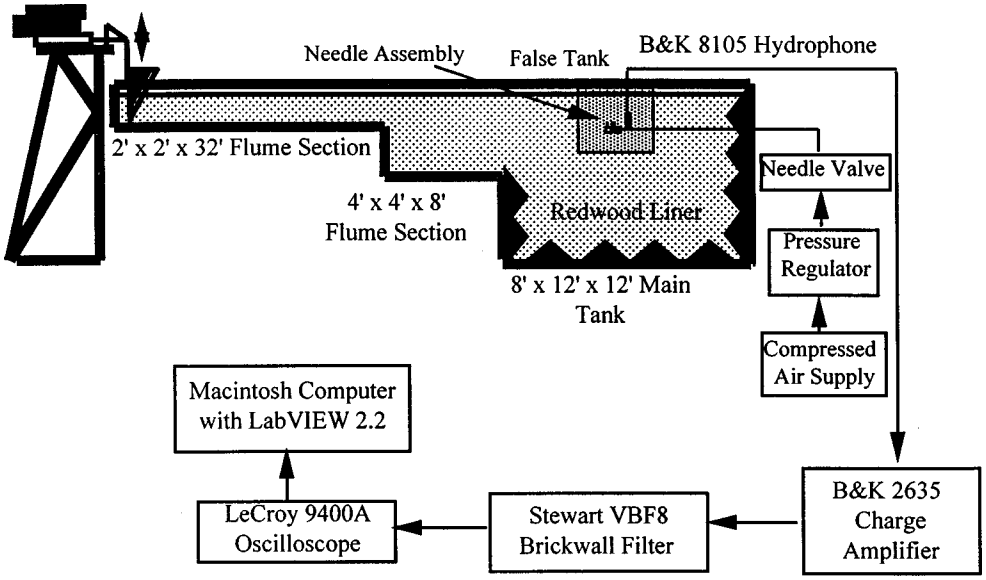


FIG. 1. Schematic drawing of experimental arrangement (not drawn to scale).

analysis. Bubbles of various sizes were generated by forcing air through needles. The false tank was originally filled with fresh tap water. The salinity of the water was changed by adding sodium chloride in increments of 0.5% (5‰). The sodium chloride was first mixed in a smaller container and then poured into the tank. The sodium chloride was chosen to isolate the effects of other chemical substances that exist in ocean water. The acoustic measurements of various size bubbles were first taken in tap water by releasing bubbles from the needle assembly, and were repeated with the same exact conditions in the water that contained various sodium chloride levels. Before taking acoustic measurements, the salt was completely diffused in the tank by stirring the water.

The surface-tension effect on the bubble detachment and subsequent acoustic radiation were also measured after adding smaller concentrations of the Triton™ 100-X, a widely used nonionic surfactant agent, in another smaller, nonanechoic tank. The Triton concentrations were added up to 25 ml [critical micelle concentration (CMC) of 6.7×10^{-3}] with 5-ml increments (0.13% by volume, or CMC of 1.35×10^{-3}). The surfactant agent was dissolved completely in the water before the acoustical data were gathered. Other mineral acids, such as HCl, which has little or no effect on bubble coalescence,[11] was used to examine the bubble sounds. In this case, the concentrations of HCl were added up to 3.0% (30‰) by volume with a 0.5% (5‰) increment. In the smaller nonanechoic tank, purified water was used.

II. RESULTS AND DISCUSSIONS

It has been reported that the presence of NaCl prevented the occurrence of bubble coalescence.[6,7] There have been numerous experimental measurements of bubble population and distribution below the sea surface. Acoustical and optical techniques have commonly been used to infer the bubble-size distribution in the upper ocean layer.[15–18] Laboratory measurements of the average bubble-size distribution generated by breaking waves in fresh and salt water in the range of 34–1200 mm in radius were given by Su *et al.*[19] and Kolaini.[1] An order-of-magnitude increase in the number density over the entire range was observed for salt water (0.3% or 30‰) versus fresh water. To demonstrate the effect of sodium chloride concentrations on the bubble population, we generated bubbles by a submerged plunging jet fitted with a cylindrical cap. The cap was hollow, with an inner diameter slightly larger than the nozzle diameter. Air was fed into the

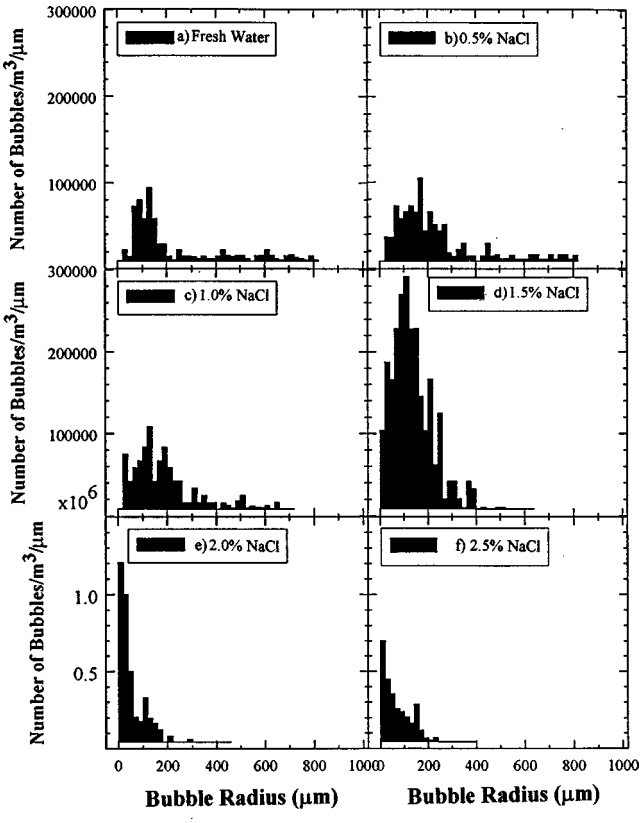


FIG. 2. Histograms of bubble population generated by a submerged water jet as a function of sodium chloride concentrations. Note the vertical scale in histograms (a)–(d) is different from that in histograms in (e) and (f).

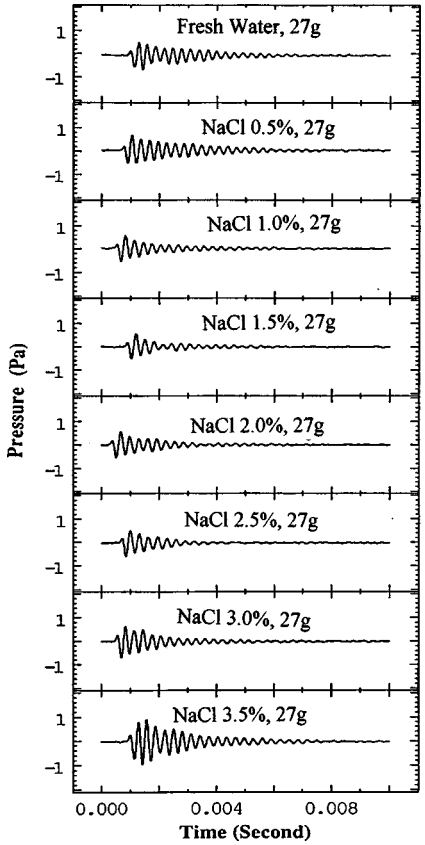


FIG. 3. The pressure–time traces of a bubble of 1.1 mm in radius pinching off from a 27-gauge needle as a function of various sodium chloride concentrations.

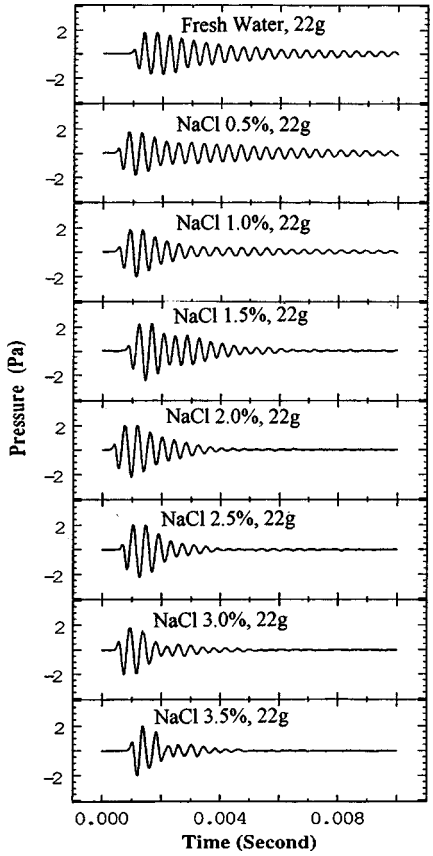


FIG. 4. The pressure–time traces of a bubble of 1.34 mm in radius pinching off from a 22-gauge needle as a function of various sodium chloride concentrations.

cap at the jet exit, and bubbles were entrained at the intersection of the jet and the liquid (see Oguz[20]). For a given jet speed, the bubble-size distributions in the tank about 3 cm below the jet exit were measured. The size distribution was measured by utilizing a high-speed video camera with a fiber-optic cable, enabling views to be obtained of individual bubbles generated by the jet. Figure 2(a)–(f) show a series of bubble-size histograms as a function of sodium chloride concentrations. Each point in these figures was averaged over two images. As expected, the bubble spectra show a shift to smaller size, and have a marked increase in density as the sodium chloride concentration increases. The smallest bubble radius measured by this optical system was 14 μm. Bubbles with a radius less than 14 μm generated in water with NaCl concentration above 2.5% (25‰) were not detected by this optical system.

Besides the effect of salt on the bubble-size populations, the sodium chloride has significant effects on the acoustical characteristics of the bubble. Figures 3–5 show the pressure–time traces of acoustic emissions from various bubble radii as a function of sodium chloride concentrations. Figure 3 is the pressure–time signature of a bubble with radius of 1.1 mm detached from a 27-gauge needle (0.19 mm in diameter). It is clear from this figure that both the pressure amplitude and quality factor of the bubble change as the sodium chloride concentrations increase. The size of the bubble remained unchanged. The striking changes in the quality factor of the bubble with sodium chloride concentra-

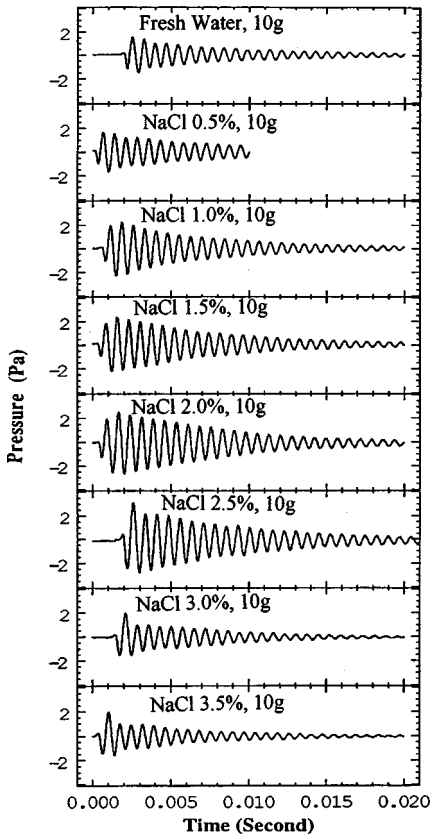


FIG. 5. The pressure–time traces of a bubble of 2.4 mm in radius pinching off from a 10-gauge needle as a function of various sodium chloride concentrations.

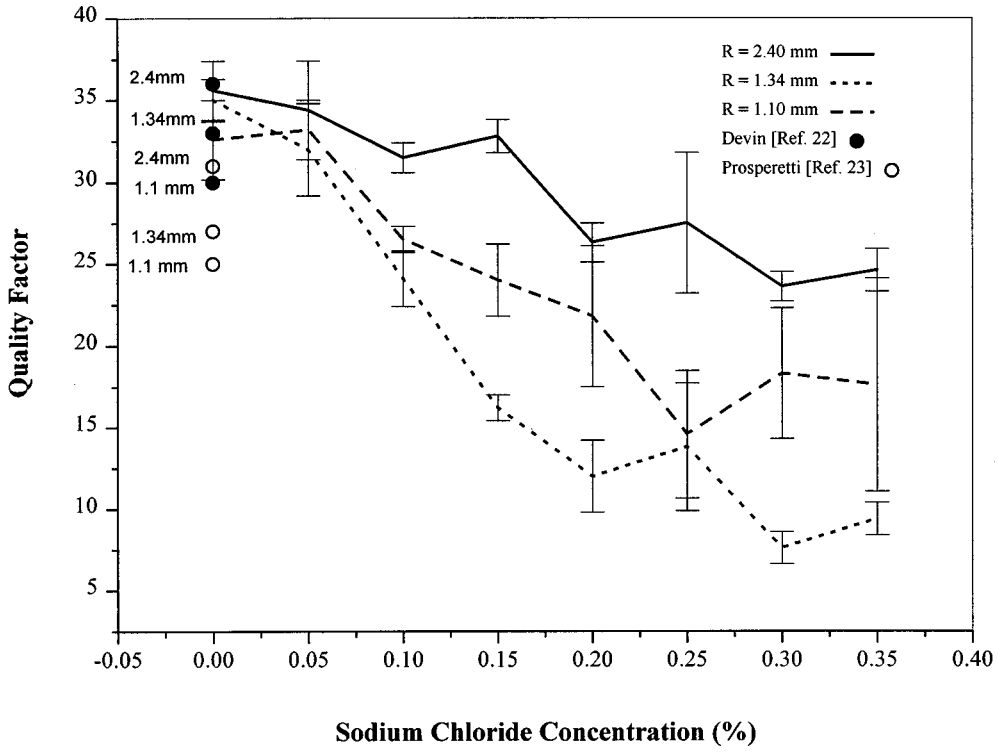


FIG. 6. The quality factor of bubbles as a function of sodium chloride concentrations. The theoretical calculations given by Devin (Ref. 22) and Prosperetti (Ref. 23) for bubble radii 1.1, 1.34, and 2.4 mm in fresh water are also shown in this figure.

tion in the range of 1.0%–3.0% (10‰–30‰) are clear. Figure 4 is the pressure–time series of a bubble with radius of 1.34 mm detaching from a 22-gauge needle (0.39 mm in diameter) in various sodium chloride concentrations. The dramatic changes in the damping coefficient of the bubble at higher salt concentrations are pronounced for this case. Finally, Fig. 5 illustrates a series of pressure–time traces of acoustic emissions from a bubble with radius of 2.4 mm detached from a 10-gauge needle (2.7 mm in diameter). In this example, the increase in sound-pressure levels as a function of salt concentrations is more pronounced than the change in quality factor of the bubble. The acoustic signatures of bubbles of various radii shown in Figs. 3–5 serve to illustrate the fact that bubble sounds are affected significantly by sodium chloride concentrations. The acoustic pressures of some of the signatures shown in Figs. 3–5, for unknown reasons, have been shown to first grow and then decay with change in sodium chloride concentrations (for example, see Fig. 3, NaCl 3.5%, and Fig. 5, NaCl of 1.0%–2.0%). It is interesting to note that the quality factor of bubbles has decreased significantly from the normal value in the fresh water to more than a third in water with 3.5% (35‰) sodium chloride, as shown in Fig. 6. It is apparent from these figures that the acoustic characteristics of the bubbles are dependent upon salinity level. These observations reveal that the effects of salt on the bubble sounds may be an important factor to consider in the understanding of the ambient sound in the ocean and bubble cavitation, and should be taken into account in the development of the bubble-dynamics model.

The first attempt to calculate the resonance frequency as a function of bubble radius was made by Minnaert.[21] The well-known Minnaert equation didn't include the damping of the bubble. The total damping δ at resonance of a bubble is given by $\delta=\delta_{ac}+\delta_{th}+\delta_{vs}$, where δ_{ac} is the acoustic damping, δ_{th} is the thermal damping, and δ_{vs} is the viscous damping. The damping coefficients were discussed in detail by Devin[22] [see Eq. (84)] and Leighton.[5] The estimated quality factor, defined as $Q=1/\delta$, in fresh water is shown in Fig. 6 for bubble radii 1.1, 1.34, and 2.4 mm. A different approach for estimating the damping coefficient was given by Prosperetti,[23] who described the thermal effects in terms of the effective polytropic index and thermal-damping constant. The estimated quality factor for these bubble radii using Prosperetti's calculation is also shown in Fig. 6. None of the present models for the damping coefficient of bubbles has included the effect of the impurity in the water.

It has been hypothesized that there are two ways in which the presence of salt can affect the bubble coalescence. One may be due to the difference in concentration of surface active components that may prevent bubble coalescence, whereas the other one was cited to be the ionic repulsive forces. To demonstrate the effect of the surfactant agents on the bubble-detachment mechanism and acoustic radiation, we mixed several concentrations of Triton 100-X with water and repeated the bubble sounds under identical conditions as before. Figure 7 shows the pressure–time traces of a 1.34-mm bubble detaching from a 22-gauge needle as a function of the surfactant-agent concentrations. It is clear from these traces that the effect of surfactant agent on the sound radiation of the bubble is not as dramatic as the effects of the sodium chloride. In this case, the size of the bubble remained the same at various Triton additives. However, there is some decrease in sound levels. The decrease in

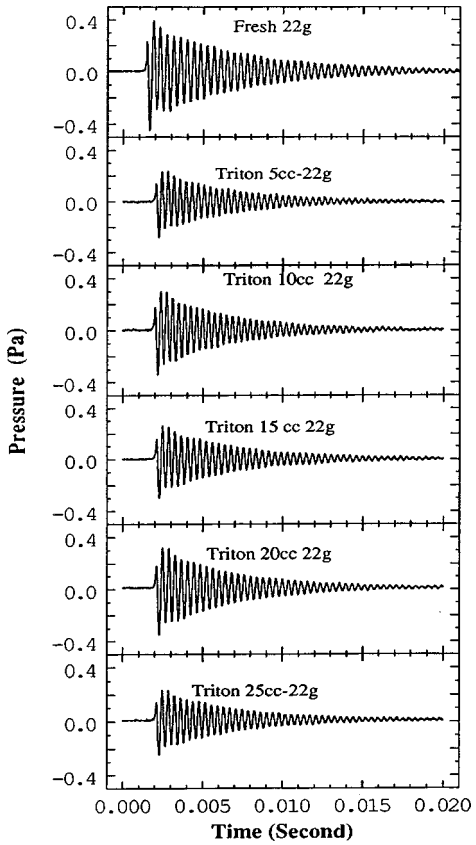


FIG. 7. The pressure–time traces of a bubble of 1.34 mm in radius pinching off from a 22-gauge needle as a function of various Triton concentrations.

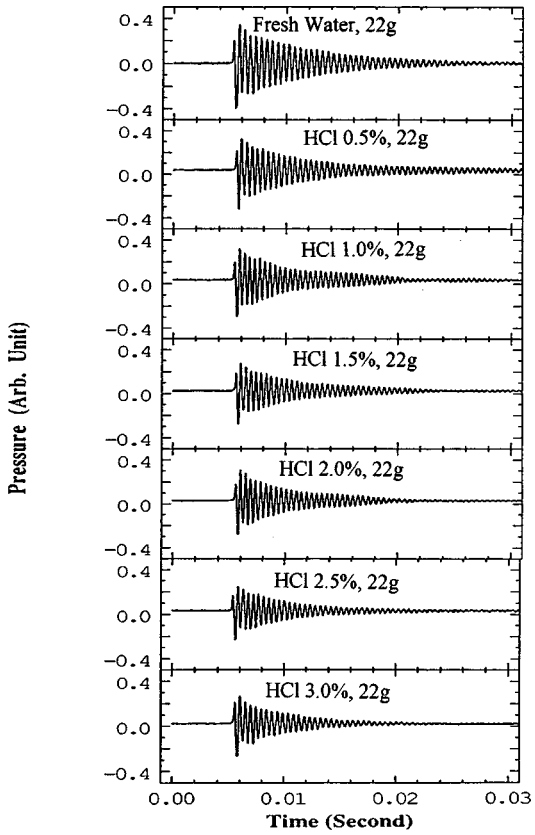


FIG. 8. The pressure–time traces of a bubble of 1.34 mm in radius pinching off from a 22-gauge needle as a function of various HCl concentrations.

sound-pressure level is more pronounced at the first Triton concentration (5 cc).

Another mineral acid, HCl, which has a minimal effect on the bubble coalescence, was also used to examine the sound generation from bubbles detaching from the needles. Figure 8 shows a series of the pressure–time traces of a 1.34-mm-size bubble detaching from a 22-gauge needle as a function of the HCl concentrations. In this solution, the acoustical behavior of the bubble remained relatively unchanged. The measured quality factor of bubbles in various HCl and Triton-100 solutions remained unchanged. It is apparent that the change in local surface tension doesn't alter the acoustic radiation of bubbles significantly. The local influence of the ions on water structure, possibly related to the ionic repulsive forces, may play a dominant role in altering the sound pressure and significantly reducing the quality factor of the bubble sound. These observations show that the effect of salt on the acoustical characteristics of bubbles is important and should be included in the models describing bubble acoustic radiation.

III. SUMMARY

It has been shown that the signature generated by laboratory breaking waves exhibits an increase in sound-pressure level. To gain some insight into the effects of salt on bubble sounds, an experiment was designed to release air from various size needles. The sounds radiated from bubbles as a function of sodium chloride concentrations were measured. The measurements show that the sound-pressure level and the damping constant of the bubbles change significantly as the sodium chloride concentrations are increased. In this study, it has been shown that the Triton additive as a surfactant agent doesn't have a significant effect on the bubble acoustic radiation. The mineral acid HCl, which has been shown to be ineffective in the coalescence of the bubble, didn't alter the acoustical characteristics of the bubble. It seems that the ionic repulsive forces present in the sodium chloride may have a direct effect on the sound radiation of bubbles. This effect is currently being investigated in greater detail.

ACKNOWLEDGMENTS

The author would like to thank Andrea Prosperetti and Hank Medwin for their valuable comments, Kathleen Stebe for her help in using the surfactant agent, Hasan Oguz for providing the water jet bubble generator, and Darin Stites, Mike Pujol, and Xinwei Hao for assisting in data acquisition. This work is supported by the Office of Naval Research.

[1] A. R. Kolaini, ''Sound radiation by various types of laboratory breaking waves in fresh and salt water,'' J. Acoust. Soc. Am. **103**, 300–308 (1998).
[2] B. R. Kerman, editor, ''Natural mechanisms of surface-generated noise in the ocean,'' in *Sea Surface Sound* (Kluwer, Dordrecht, The Netherlands, 1988).
[3] B. R. Kerman, editor, ''Natural mechanisms of surface-generated noise in the ocean,'' in *Natural Physical Sources of Underwater Sound* (Kluwer, Dordrecht, The Netherlands, 1993).
[4] M. J. Buckingham and J. R. Potter, ''Sea Surface Sound 94,'' *Proceed-*

ings of the Third International Meeting on Natural Physical Sources of Underwater Sound (World Scientific, New Jersey, 1995).
[5] T. G. Leighton, *The Acoustics of Bubbles* (Academic, New York, 1994).
[6] J. C. Scott, ''The role of salt in whitecap persistence,'' Deep-Sea Res. Oceanogr. Abstr. **22**, 653–657 (1975).
[7] E. C. Monahan and C. R. Zeitlow, ''Laboratory comparisons of fresh and salt water whitecap,'' J. Geophys. Res. **74**, 6961–6966 (1969)
[8] J. A. Kitchner, ''Foams and free liquid film,'' in *Recent Progress in Surface Science*, Vol. 1 (Academic, New York, 1964), pp. 51–93.
[9] C. Pounder, ''Sodium chloride and water temperature effects on bubbles,'' in *Oceanic Whitecaps and Their Role in Air-Sea Exchange Processes*, edited by E. C. Monahan and M. N. Georoid (Reidel, Boston, 1986).
[10] V. I. Klassen and V. A. Mokrousov, *An Introduction to the Theory of Flotation* (Butterworths, London, 1963).
[11] V. S. J. Craig, B. W. Ninham, and R. M. Pashley, ''The effect of electrolytes on bubble coalescence in water,'' J. Phys. Chem. **97**, 10192–10197 (1993).
[12] W. M. Carey, J. W. Fitzgerald, E. C. Monahan, and Q. Wang, ''Measurements of the sound produced by a tipping trough with fresh and salt water,'' J. Acoust. Soc. Am. **93**, 3178–3192 (1993).
[13] A. R. Kolaini, R. A. Roy, and D. C. Gardner, ''Low-frequency acoustic emission in fresh and salt water,'' J. Acoust. Soc. Am. **96**, 1766–1772 (1994).
[14] A. Prosperetti, ''Bubble dynamics in oceanic noise,'' in *Sea Surface Sound: Natural Mechanisms of Surface Generated Noise in the Ocean*, edited by B. R. Kerman (Kluwer Academic, Boston, 1988), pp. 151–171.
[15] H. Medwin and N. D. Breitz, ''Ambient and transient bubble spectral densities in quiescent areas and under spilling breakers,'' J. Geophys. Res. **84**, 12751–12759 (1989).
[16] P. A. Kolovayev, ''Investigation of the concentration and statistical size distribution of wind-produced bubbles in the near-surface ocean layer,'' Oceanology **15**, 1013–1017 (1975).
[17] B. D. Johnson and R. C. Cooke, ''Bubble populations and spectra in coastal waters: A photographic approach,'' J. Geophys. Res. **84**, 3761–3766 (1979).
[18] J. Wu, ''Bubbles in the near-surface ocean: a general description,'' J. Geophys. Res. **93**, 587–590 (1987).
[19] M. Y. Su, D. Todoroff, and J. Cartmill, ''Laboratory comparisons of acoustic and optical sensors for microbubble measurements,'' J. Atomos. Ocean. Technol. **11**(1), 170–181 (1994).
[20] H. N. Oguz, ''The role of surface disturbances in the entrainment of bubbles by a liquid jet,'' J. Fluid Mech. **372**, 189–212 (19989).
[21] M. Minnaert, ''On musical air-bubbles and the sounds of running water,'' Philos. Mag. **16**, 235–248 (1933).
[22] C. Devin, ''Survey of thermal, radiation, and viscous damping of pulsating air bubbles in water,'' J. Acoust. Soc. Am. **31**, 1654–1667 (1959).
[23] A. Prosperetti, ''Thermal effects and damping mechanisms in the forced radial oscillations of gas bubbles in liquids,'' J. Acoust. Soc. Am. **61**, 17–27 (1977).

The role of scale structure in scattering from random rough surfaces

Vincent Lupien
Department of Ocean Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 18 July 1997; accepted for publication 18 November 1998)

Scale structure is defined as the stochastic spatial arrangement of an interface's component features at various scales. The principal contribution of this study is to demonstrate that scale structure is a useful and acoustically meaningful concept for characterizing a stochastic scattering surface's roughness. The information contained in scale structure about an interface's roughness supplements that contained in the power spectral density (PSD). This fact is demonstrated by observing that the scale structure of different stochastic surface models with a common power spectral density of the power-law type varies from feature-like at one extreme to fractal, non-feature-like at the other. To investigate the impact of scale structure in scattering, the ensemble scattering statistics of each surface model are computed using an exact Monte Carlo integral equation technique and 50 realizations each. It is found that both the bistatic scattering strength and the time-domain statistics of scatter are significantly affected by differences in scale structure given a constant PSD. This result is relevant in any scenario involving the scattering of waves from natural interfaces because while landforms usually have a power-law PSD, they exhibit a wide range of scale strucutres. To contrast the roles played by scale structure and the PSD in scattering, surface models with vastly *different* PSDs but similar scale structure are compared. It is found that while the PSD's role is preeminent for average scattering strength in the back quadrant, scale structure establishes the presence of a forward scatter lobe and the degree of event-like (non-Rayleigh) behavior in time-domain scatter; both are enhanced in feature-like surfaces. As a new way of looking at roughness, scale structure suggests many possibilities for future research, some of which are presented herein. © *1999 Acoustical Society of America.* [S0001-4966(99)02703-4]

PACS numbers: 43.30.Hw, 43.20.Fn [DLB]

INTRODUCTION

A. Motivation

The work contained in this paper was motivated by analysis of acoustical and bathymetric data from the Mid-Atlantic Ridge in Ref. 1. The power spectral estimates of profiles collected normal and parallel to the axis of anisotropy of a particular bathymetric region obey a power-law relationship. Interestingly, the spectra are quasi-identical in both shape and magnitude despite visually obvious differences in the structure of the roughness normal and parallel to the anisotropy. While both directions exhibit *feature-like* roughness, the scale and shape of the features are different. Concurrently, both the backscattering strength and the probability density functions (pdf's) of backscattered envelopes are distinct along the two directions. These observations have suggested (i) that the power spectral density is not always sufficient to differentiate interfaces with obvious differences in roughness and (ii) that the component of an interface's morphology which is left unspecified by the second moment is relevant to acoustic scattering. A method for more fully characterizing the roughness is needed.

B. The problem of characterizing roughness

It is clear that an interface's *roughness* and the material properties on either side of it govern the scattering process. It is not a trivial matter to decide on a characterization of the roughness which, on the one hand, avoids pedantic exactness while, on the other, embodies all the relevant aspects of the roughness for the desired application.

To illustrate this fundamental trade-off, consider an example from the very beginning of the field of scattering from rough surfaces. In the late 19th century, Lord Rayleigh considered scattering of a plane monochromatic wave with wavenumber k incident at a grazing angle of θ_i on a sinusoidal surface with amplitude Δh. The roughness was described by the single parameter $R_a = k\Delta h \cos\theta_i$, which is now known as the *Rayleigh* parameter.[2] A descriptor of surface roughness such as the Rayleigh parameter has two components: an ''interface'' component describing the interface corrugations themselves (in this case Δh) and an acoustic component involving parameters of the incident field that imposes a relative scale on the interface component (in this case $k \cos\theta_i$). The focus in this article is on the interface component.

In a statistically rough surface, the interface component Δh is replaced by the root-mean-square (rms) height of the surface. Using only this limited information about the surface itself, it is possible to explain the loss of coherence in the specular direction as the surface amplitude increases relative to the wavelength, demonstrating the power of simple but intelligently chosen characterizations. On the other hand, predictions of scattering based solely on the Rayleigh parameter are inadequate for most modern applications.

At the other extreme, one could attempt to achieve a

complete probabilistic description of a surface including all statistical moments and any possible nonstationarity, an approach which might give perfect results if only it were not utterly impractical and, in fact, impossible. Physical insight must be used to devise an optimal model which characterizes all that is relevant in a given scattering scenario while omitting what is *ir*relevant.

For example, to satisfy the modern need for greater accuracy, the Rayleigh parameter is usually supplanted by expressions involving the full correlation function of a surface, which has the desirable property of characterizing both the horizontal and vertical scales. While the correlation function is unarguably superior to the rms height alone, it is shown here that it is insufficient for the purposes of calculating bistatic scattering strength within 6 dB, and that it is especially inadequate in applications involving the time-domain scattering of high-resolution pulses from interfaces having a power-law power spectral density (PSD).

A natural impulse to resolve this inadequacy is to extend formulations of scatter to include the higher-order statistical moments of the surface. A good example is the Kirchhoff approximation, which in its most common form involves only the second moment but in its general form involves all of the surface's moments. While including more moments is mathematically sound, a decreasing portion of the information contained in each higher moment is physically meaningful to scattering. A large number of moments may therefore be required to obtain the desired improvement, greatly complicating analytical expressions and pushing computational requirements to excessive levels. Further, while the second moment is physically intuitive, the information in each higher moment is increasingly obscure to interpret.

In this paper, a new concept for characterizing surface roughness is introduced. It is called *scale structure*, and is defined as the stochastic spatial arrangement of the component features of an interface at various scales. The numerical experiments presented here demonstrate that the scale structure of an interface provides physically meaningful insight into scattering properties, beyond what the second moment can achieve, and at a level of intuitiveness and simplicity that is hard to imagine from higher moment expansions. These statements are tempered by the qualitative nature of scale structure, but the wavelet transform is shown to be a promising tool for a quantitative definition. While the quantitative approach sketched out in this paper does require that moments higher than the second be considered, it is anticipated that it will maintain a high level of intuitiveness because it uses the statistics of wavelet coefficients, which are directly linked to scale structure, instead of the statistics of the surface elevations.

C. Approach

The metrics through which the effects of scale structure in scattering are assessed are (i) the ensemble-averaged narrowband bistatic scattering strength and (ii) the pdf's of wideband time-domain backscattered envelopes. The incident field is a tapered plane wave at a grazing angle of 45 degrees and the surfaces are one-dimensional rigid profiles.

Four stochastic surface models are used in the study. Three of them have the same PSD (of the power-law type) but different scale structure ranging from feature-like at one extreme to maximally non-feature-like (fractal) at the other. In the first numerical experiment, the ensemble statistics of scatter are obtained for the three models having power-law PSD at an rms height and correlation length of $(\sigma, l_c) = (0.3, 4.0)\lambda$, where λ is the acoustic wavelength. The scatter statistics for each model are computed using a Monte Carlo method based on an exact integral equation solution.[4] The results show that, given a common PSD, surface models with different scale structure have different scattering properties.

In the second numerical experiment, the fourth surface model, which has a Gaussian PSD, is compared to the feature-like and non-feature-like models with power-law PSD. The Gaussian PSD has a much more rapidly decaying tail at high wavenumbers than the power-law PSD, thereby yielding roughness which is always feature-like. Interestingly, scatter from the surface with Gaussian PSD is found to share some of the characteristics of scatter from the feature-like model with power-law PSD, indicating that scale structure prescribes certain scattering properties independently of the PSD. Here, the values of rms height and correlation length studied are $(\sigma, l_c) = (0.3, 4.0)\lambda$ and $(0.2, 1.0)\lambda$.

All the results are presented in nondimensionalized form and are independent of center frequency and sound speed; they are thus applicable in a wide range of physical settings.

The four surface models are discussed in Sec. I. In Sec. II we present the theory behind the integral equation solution, Sec. III contains the results of the numerical experiments, and, finally, in Sec. IV we summarize the important findings. First, the next subsection presents some related previous work.

D. Previous work

In Ref. 3, the authors study whether surfaces that are different in shape yet have the same power spectrum scatter energy in the same manner. In the first study, a sawtooth surface is generated; its power spectrum is calculated and three new surfaces are generated by randomly scrambling the phases at each wavenumber. The three scrambled-phase surfaces are less feature-like than the sawtooth surface but have the same energy at each wavenumber. The scattering strength and coherence are found to be distinct for all four surfaces, yielding the important conclusion that differences in shape given a common power spectrum *do* affect scattering properties. In the second study, scatter from a water-wave realization is compared to scatter from three realizations having identical spectra but scrambled phases, yielding identical conclusions to the first study.

Instead of beginning with an individual surface and perturbing its phase components, the present paper begins with four stochastic surface models each of which describes the ensemble statistics of a *family* of surfaces. The focus is on the effect of a change in the ensemble statistics of the surface models on the ensemble statistics of scatter. To quantify this effect, 50 realizations from each model are generated (whose

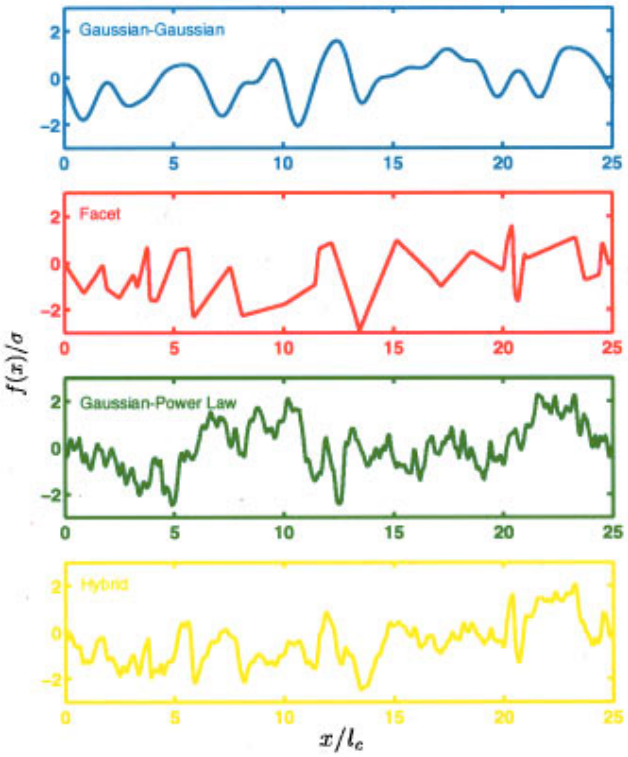


FIG. 1. Sample paths for the GG, facet, GPL, and hybrid processes, normalized to unit rms height and correlation length. The GG and facet realizations exhibit feature-like scale structure while the GPL and hybrid realizations appear non-feature-like. The hybrid realization is a weighted sum of the facet and GPL realizations with 75% of the energy coming from the facet realization, demonstrating that a small amount of non-feature-like noise can mask a feature-like signal. The facet, GPL, and hybrid processes have the same PSD, demonstrating the insufficiency of the second moment in fixing scale structure.

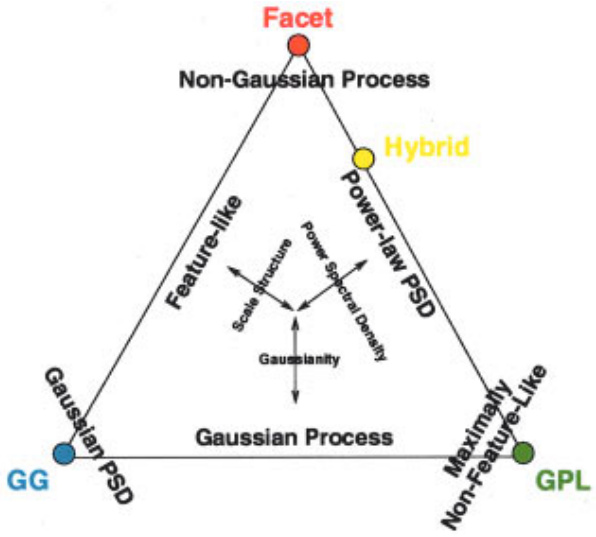


FIG. 2. The facet, GPL, and GG surfaces define the three poles of a surface property triangle. The three arrows in the center define the axes along which surface properties can vary. The GG surface is Gaussian, has Gaussian PSD, and has feature-like roughness. The GPL surface is Gaussian, has power-law PSD, and is maximally non-feature-like. The facet surface is non-Gaussian, has power-law PSD, and has feature-like roughness. The hybrid surface lies along a constant value of the PSD axis between the GPL and facet poles.

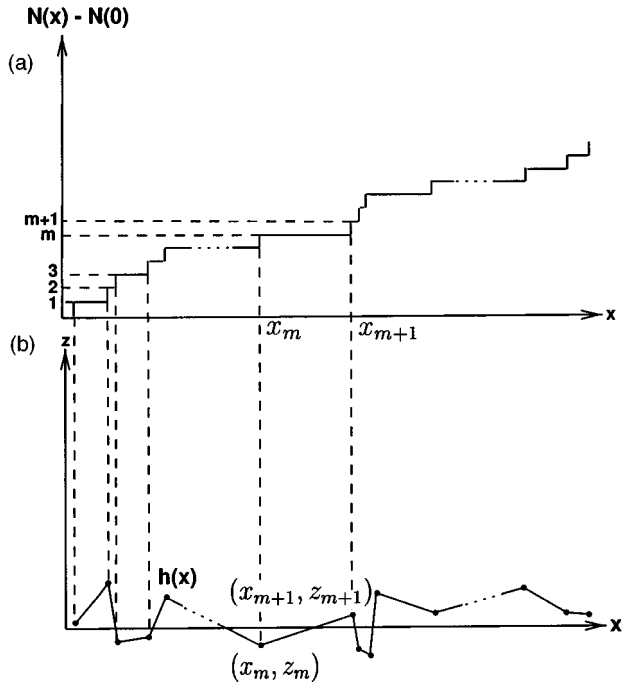


FIG. 3. (a) A Poisson counting process $N(x)$ defines a set of transition locations x_i. (b) The facet process $h(x)$ is constructed by joining adjacent vertices (x_i, z_i) with straight line segments where the z_i are uncorrelated, zero-mean, and Gaussian.

Fourier transforms differ in both amplitude and phase), from which scatter is computed exactly and ensemble averaged. The chosen focus enables a conceptual study that aims to determine which aspects of stochastic roughness are relevant to scattering and how they can be effectively and efficiently described.

The emphasis on stochastic models and their resulting ensemble scattering properties is inspired from Refs. 4–6, as is the Monte Carlo methodology based on an exact integral

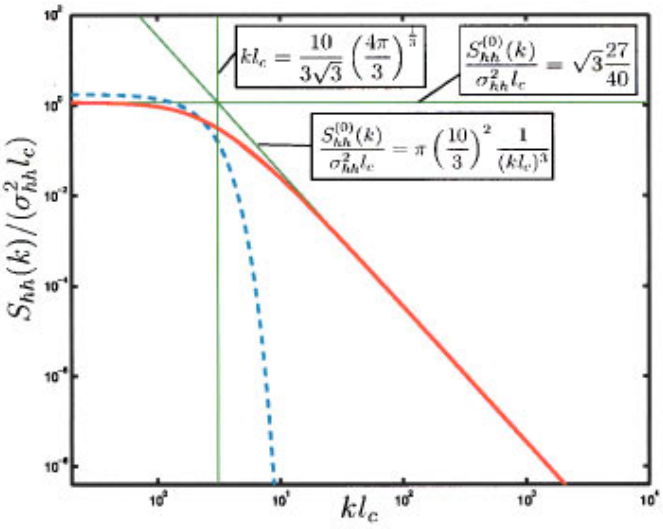


FIG. 4. Normalized facet spectrum in log–log scale (red), representing the PSD of the facet, GPL, and hybrid processes. The corner wavenumber along with the high- and low-frequency asymptotes are shown in green. A normalized Gaussian spectrum, representing the PSD of the GG process, is shown in blue.

equation formulation. In Refs. 4–6, exact average bistatic scattering strengths from pressure-release Gaussian surfaces with Gaussian and Pierson–Moskowitz spectra are computed. The present paper concentrates on identifying the role played by scale structure versus the role played by the PSD. The purpose of Refs. 4–6 was different, namely to investigate the validity of the perturbation and Kirchhoff approximations. However, the two models studied appear to be good ones for the purposes of the present article because they differ in PSD and scale structure. The Pierson–Moskowitz model has a power-law PSD and non-feature-like scale structure, while the surface with Gaussian PSD has feature-like scale structure. Since the computations for the two models in Refs. 4–6 were not performed at consistent values of rms height and correlation length, comparison of their scattering properties is difficult. Additionally, because the two models differed in *both* scale structure and PSD, it is impossible to distinguish which parameter is responsible for the observed differences.

Here, a new surface model is introduced which has a power-law PSD yet has feature-like scale structure. The roles of scale structure and the PSD are then systematically studied by comparing the properties of the new model (the facet model) to those of a non-feature-like model with identical power-law PSD (the Gaussian–power–law model), a feature-like model with Gaussian PSD (the Gaussian–Gaussian model), and a model which is a linear combination of the facet and Gaussian–power-law models (the hybrid model). The surface models are described in detail in the following section. The pdf's of *time-domain* backscatter are computed along with the bistatic scattering strengths. Each numerical experiment is performed at identical values of rms height and correlation length for each surface model.

I. STOCHASTIC SURFACE MODELS

This section provides an in-depth look at the four surface models used in the study: Gaussian–Gaussian (GG), facet, Gaussian–power-law (GPL), and hybrid. A sample realization of each of these models is depicted in Fig. 1, normalized to unit variance and correlation length. The first two surface models share the quality of being feature-like in spite of having vastly different PSDs. Conversely, the facet, GPL, and hybrid processes exhibit large differences in scale structure in spite of having exactly the same PSD.

In Fig. 2, the properties of the four surface models are represented graphically. The GG, facet, and GPL models lie at the poles of a surface property triangle with axes corresponding to Gaussianity, PSD and scale structure.

A. The Gaussian–Gaussian (GG) surface

The GG surface has Gaussian statistics and a Gaussian power spectral density. Since it is possible to get confused between Gaussian statistics and a Gaussian PSD, this section provides a definition for both to establish that they are in no way connected and that they make quite different statements about a random process.

The complete specification of a one-dimensional random process $f(x)$ requires the N-point joint probability density of heights $p_{\mathbf{f}}(\mathbf{F})$, where $\mathbf{f}=[f_1,f_2,\ldots,f_N]^T=[f(x_1),f(x_2),\ldots,f(x_N)]^T$ is a column vector of heights at arbitrary locations $\mathbf{x}=[x_1,x_2,\ldots,x_N]^T$ for any N. Specification of the mean $\mu(x)=\mathcal{E}[f(x)]$ and correlation function $R_{ff}(x',x'')=\mathcal{E}[f(x')f(x'')]$ fixes the first and second moments of $p_{\mathbf{f}}(\mathbf{F})$:

$$\boldsymbol{\mu}=\int \mathbf{F}\,p_{\mathbf{f}}(\mathbf{F})\,d\mathbf{F} \tag{1}$$

and

$$\mathbf{R}=\int \mathbf{F}\mathbf{F}^T p_{\mathbf{f}}(\mathbf{F})\,d\mathbf{F}, \tag{2}$$

where $[\boldsymbol{\mu}]_i=\mu(x_i)$ and $[\mathbf{R}]_{ij}=R_{ff}(x_i,x_j)$. When only the first and second moments are specified, the expectations of products of three or more (possibly repeated) elements of $\mathbf{f}$ are unknown. In particular, none of the moments greater than two for even the *one*-point probability density $p_{f_i}(F_i)$ are known. The unspecified moments leave room for enormous variability which can leave scale structure undetermined.

All the stochastic processes used in this paper for surface modeling are zero-mean and wide-sense stationary. The latter condition implies that $R_{ff}(x',x'')=R_{ff}(\chi)$, where $\chi=|x''-x'|$.[7] The power spectral density $S_{ff}(k)$ is then defined as the Fourier transform of $R_{ff}(\chi)$:

$$S_{ff}(k)=\int_{-\infty}^{\infty} R_{ff}(\chi)e^{-ik\chi}\,d\chi, \tag{3}$$

where k is the wavenumber.

A random process is *Gaussian* if, for all sample locations $\{x_i\}$ and arbitrary dimensionality N,[8]

$$p_{\mathbf{f}}(\mathbf{F})=\frac{1}{(2\pi)^{N/2}|\boldsymbol{\Lambda}|^{1/2}}\exp\left(-\frac{1}{2}\mathbf{F}^T\boldsymbol{\Lambda}^{-1}\mathbf{F}\right), \tag{4}$$

where $\boldsymbol{\Lambda}=\mathcal{E}[(\mathbf{f}-\boldsymbol{\mu})(\mathbf{f}-\boldsymbol{\mu})^T]$, and in this case $\boldsymbol{\Lambda}=\mathbf{R}$ since $\boldsymbol{\mu}=\mathbf{0}$. In wide-sense stationary Gaussian processes, specification of the mean and correlation function is a complete probabilistic description. The first and second moments determine all higher-order moments in the N-point density.

The term ''Gaussian process'' is used when Eq. (4) applies. On the other hand, a process has Gaussian PSD when $R_{ff}(\chi)$ [or, equivalently, $S_{ff}(k)$] has a Gaussian shape, as in

$$\frac{S_{ff}(k)}{\sigma_{ff}^2 l_c}=\sqrt{\pi}e^{-(kl_c)^2/4}, \tag{5}$$

where l_c is the correlation length and σ_{ff}^2 is the variance. Neither description implies the other.

The GG surface is both a Gaussian process and has a Gaussian power spectral density. It has feature-like scale structure and occupies the lower right pole in Fig. 2.

B. The facet surface

The facet surface is constructed so as to be feature-like by joining random vertices $\{(x_i,z_i)\}$ with straight line segments through the equation

$$h(x)=z_m+\left(\frac{x-x_m}{x_{m+1}-x_m}\right)(z_{m+1}-z_m),\quad x_m\leqslant x<x_{m+1}, \tag{6}$$

where the $\{x_i\}$ are Poisson arrivals and the $\{z_i\}$ are zero-mean, Gaussian and independent of the $\{x_i\}$.

The Poisson arrivals $\{x_i\}$ are the transitions of the following Poisson counting process $N(x)$:

$$Pr[N(x'')-N(x')=l]=\frac{[\lambda_r(x''-x')]^l}{l!}e^{-\lambda_r(x''-x')}, \tag{7}$$

where $N(x'')-N(x')$ is the count of the number of transitions between x' and x''. It describes a staircase in which the length of each step is random with mean $1/\lambda_r$ and the height of each step is 1, as shown in Fig. 3(a). The $\{x_i\}$ are defined as the locations of the transitions of $N(x)$. The larger λ_r, the more quickly new arrivals occur.

The Gaussian heights $\{z_i\}$ are zero-mean with variance σ_{zz}^2, having the probability density

$$p_{z_i}(Z_i)=\frac{1}{\sqrt{2\pi}\sigma_{zz}}e^{-z_i^2/2\sigma_{zz}^2}, \tag{8}$$

and stationary discrete correlation function $R_{zz}[l]=\mathcal{E}[z_jz_i]$, where $\mathcal{E}$ is the expectation operator and $l=|j-i|$.

Let $h^{(0)}(x)$ be the case of $h(x)$ where the set $\{z_i\}$ is a discrete Gaussian white noise sequence:

$$R_{zz}^{(0)}[l]=\begin{cases}\sigma_{zz}^2, & l=0,\\ 0, & l\neq 0.\end{cases} \tag{9}$$

A sample realization for this case is shown in red in Fig. 1. The facet process so defined has the following correlation function:[1]

$$\begin{aligned}R_{hh}^{(0)}(\chi)&=\mathcal{E}[h(x')h(x'')]\\&=\mathcal{E}\left[\left(z_m+\left(\frac{x'-x_m}{x_{m+1}-x_m}\right)(z_{m+1}-z_m)\right)\right.\\&\quad\left.\times\left(z_n+\left(\frac{x''-x_n}{x_{n+1}-x_n}\right)(z_{n+1}-z_n)\right)\right]\\&=\sigma_{zz}^2(2E_4(\lambda_r\chi)+E_2(\lambda_r\chi)*E_2(\lambda_r\chi)),\end{aligned} \tag{10}$$

where $*$ denotes convolution, $x_m\leqslant x'<x_{m+1}$, $x_n\leqslant x''<x_{n+1}$ and $\chi=|x''-x'|$. $E_n(x)$ is the nth-order exponential integral for $x\geqslant 0$ and is zero for $x<0$:

$$E_n(x)=\begin{cases}\displaystyle\int_1^\infty e^{-xu}u^{-n}\,du, & x\geqslant 0,\\ 0, & x<0.\end{cases} \tag{11}$$

The variance of $h^{(0)}(x)$ is $(\sigma_{hh}^{(0)})^2=\frac{2}{3}\sigma_{zz}^2$, and its correlation length, calculated according to the definition in Appendix A, is

$$l_c^{(0)}=\frac{1}{\lambda_r}\frac{10}{3\sqrt{3}}. \tag{12}$$

The power spectral density of $h^{(0)}(x)$ is

$$\begin{aligned}S_{hh}^{(0)}(k)=\frac{2\sigma_{zz}^2}{\lambda_r}\Bigg[&\frac{2\arctan(k/\lambda_r)}{(k/\lambda_r)^3}\\&+\frac{1}{(k/\lambda_r)^4}\Big(\ln^2(\sqrt{1+(k/\lambda_r)^2})\\&-\ln(1+(k/\lambda_r)^2)-\arctan^2(k/\lambda_r)\Big)\Bigg].\end{aligned} \tag{13}$$

This spectrum is *power-law* because of its asymptotic behavior at large k,

$$S_{hh}^{(0)}(k)\to 2\pi\frac{\sigma_{zz}^2\lambda_r^2}{k^3}=\pi\left(\frac{10\sigma_{hh}^{(0)}}{3l_c^{(0)}}\right)^2\frac{1}{k^3}\quad\text{as } k\to\infty, \tag{14}$$

which can be observed from the straight line high-wavenumber asymptote of the log–log plot of Fig. 4. The facet surface has feature-like scale structure, as is evident in the realization of Fig. 1, and it occupies the top pole in the surface property triangle of Fig. 2, having non-Gaussian statistics.

The spectrum of the facet process is similar in shape to the estimated spectra of many natural processes, ranging from seafloor shapes to water waves and the velocity profiles of turbulent flows. Like the facet process, some of these processes exhibit feature-like roughness. However, unlike the facet process, natural processes with power-law spectra usually exhibit roughness at many different scales, giving them, at least qualitatively, a resemblance to fractal objects. Since processes with power-law PSD and Gaussian statistics are fractal in the mathematical sense,[9] they are commonly used as models. Unfortunately, they cannot capture feature-like roughness as they result from superimposed roughness contributions at all scales *continuously*; the cumulative effect of the contributions at adjacent scales is to mar the appearance of feature-like roughness.

The reason that natural interfaces with power-law PSD can exhibit roughness over such a wide breadth of scales yet still retain a feature-like appearance is that they are formed by the superposition of individual feature-like contributions acting at *discrete, widely separated* scales. The facet process appears to be a promising building block to simulate such interfaces. In Ref. 1, a composite seafloor model that is the superposition of facet processes at different scales has been used to simulate the roughness of Mid-Atlantic Ridge seafloor.

C. The Gaussian power-law (GPL) surface

Gaussian processes with power-law PSD represent a maximal departure from feature-like surfaces; the component features are superimposed at a continuum of scales and at independent locations spatially resulting in wavelet transform coefficients which are independent across both scale and space.[9] This extreme form of non-feature-like roughness is referred to in this paper as *maximally non-feature-like*.

The Gaussian–power-law process is defined as the unique process which has the PSD of the facet process and is maximally non-feature-like, i.e., Gaussian. The unique Gaussian process $f(x)$ which has the PSD of the facet process $S_{hh}^{(0)}(k)$ can be obtained from the convolution

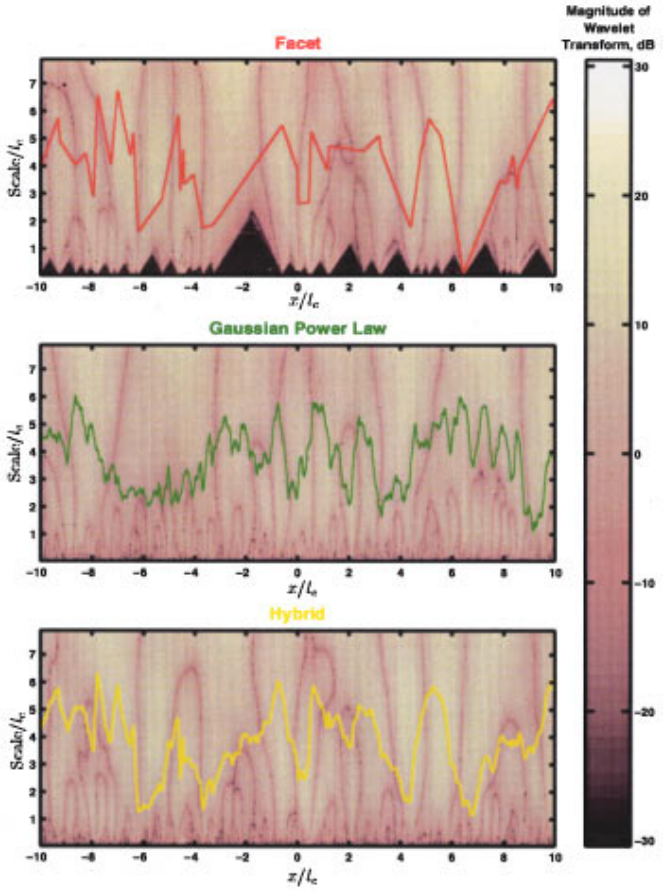


FIG. 5. Contour plot of the magnitude of the continuous wavelet transform (CWT) using a Daubechies 2 wavelet, in dB. The x axis is spatial position along the surface and the y axis is the scale parameter corresponding to the spatial extent of the analyzing wavelet. At the top of each figure, the analyzing wavelet is eight correlation lengths long, and at the bottom, its length is infinitesimal. The CWT does not significantly enhance one's ability to distinguish between the hybrid (yellow) and GPL (green) realizations. The dark areas in the facet case (red) occur when the size of the wavelet is smaller than the horizontal extent of a facet.

$$f(x)=r_{hh}^{(0)}(x)*w(x), \tag{15}$$

where

$$r_{hh}^{(0)}(x)=\mathcal{F}^{-1}\{|[S_{hh}^{(0)}(k)]^{1/2}|\}, \tag{16}$$

$w(x)$ is a zero-mean stationary white Gaussian noise process with unit variance and $\mathcal{F}^{-1}$ is the inverse Fourier transform operator.

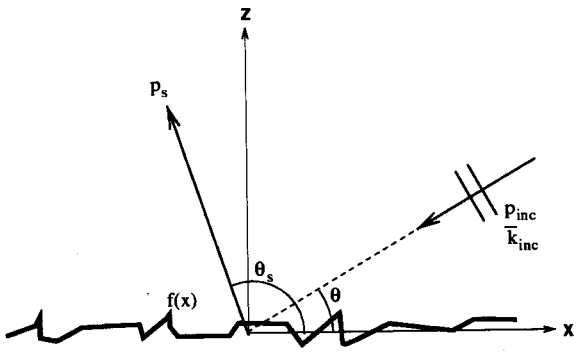


FIG. 7. Scattering geometry.

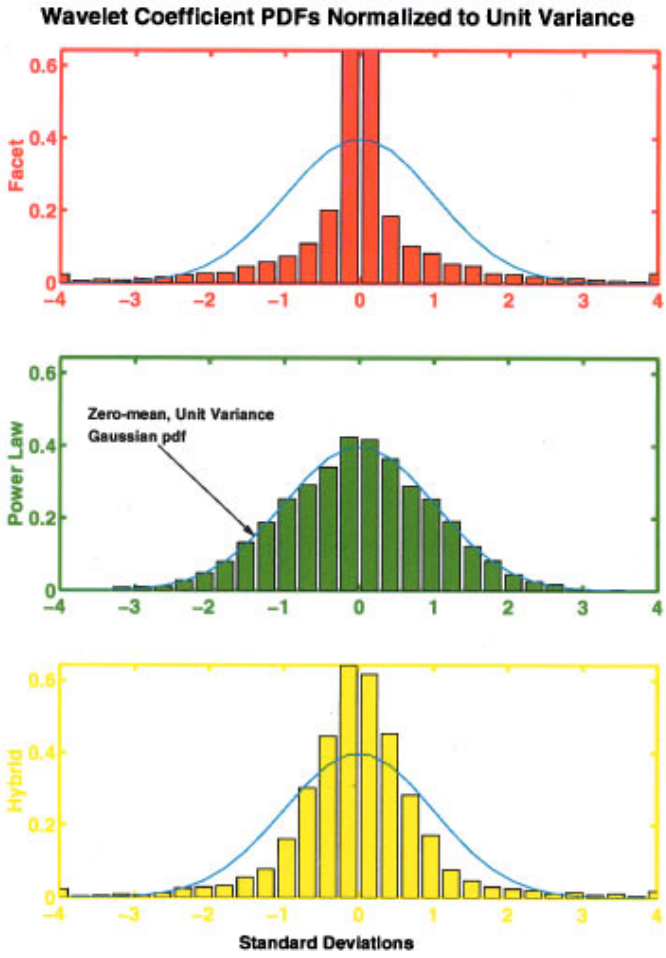


FIG. 6. Estimated pdf's of Daubechies 2 wavelet coefficients at a scale of $\frac{1}{10}$ of a correlation length, normalized to unit variance. The blue curves are unit variance Gaussian pdf's. The GPL process (green) has Gaussian coefficients, but the other two processes yield non-Gaussian coefficients. The statistical wavelet analysis has successfully differentiated the scale structure of the three processes.

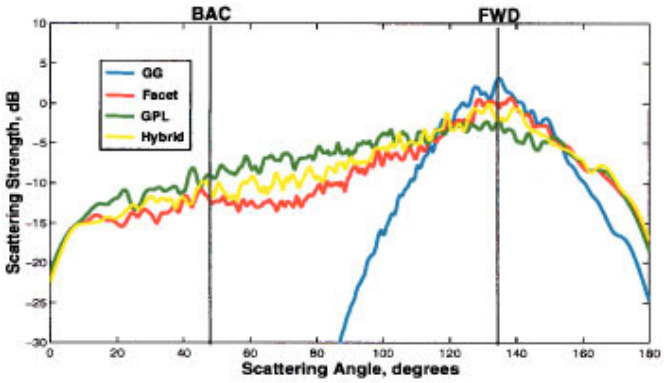


FIG. 8. Monte Carlo estimates of scattering strength for $(\sigma,l_c)=(0.3,4.0)\lambda$. The facet, GPL, and hybrid surfaces, which share a common PSD, have distinct scattering strengths emphasizing the importance of scale structure. The feature-like roughness of the facet and GG surfaces yields an enhanced forward scatter lobe compared to the GPL surface. The backscatter in the power-law models is enhanced compared to the GG surface. In all cases, energy conservation is satisfied to within 1%.

The GPL process occupies the bottom right pole in Fig. 2; a realization is shown in green in Fig. 1. Comparing it with the realization of the Facet process, which has identical correlation function, it can be seen that there are immense differences in scale structure.

The observed differences in morphology are not specific to the individual realizations shown in Fig. 1. They apply consistently to any set of realizations from the two processes and result from the fact that in the GPL surface, the roughness contributions are evenly distributed in space at all wavenumbers while in the facet process, the high-wavenumber roughness always combines in organized ways to give a combination of sharp corners and flat segments. Comparison of the facet and GPL processes clearly demonstrates that *the second moment is insufficient to specify scale structure*.

The discussion now turns briefly to two existing terms for referring to scale structure, which are ''multi-scale'' and ''single-scale.''[4,6] A fitting term for describing the scale structure of the GPL surface is ''multi-scale,'' because it presents features ranging continuously in size from the scale of the rms height and correlation length down to infinitesimal scales. The term ''single-scale'' applies excellently to the scale structure of the GG surface which shows smooth, rounded features whose dimensions are closely clustered about a mean value determined by the variance and correlation length. The facet surface is not easily categorized because it presents smooth features (facets) of nonvanishing dimension like the GG surface but has a much broader distribution in the size of its features. It appears that ''single-scale'' is more fitting that ''multi-scale.'' The misguided practice of using the two terms as descriptors of the PSD would rank the facet surface as multi-scale because of its power-law form, which is clearly not an accurate description. Proper usage of these terms refers to roughness characteristics that go beyond the PSD. Their current use reflects a nascent recognition of the physical relevance of scale structure. In this paper roughness is described through the degree of ''feature-like'' scale structure and by referring explicitly to the PSD when needed.

D. The hybrid surface

The hybrid process is created by an energy-conserving linear combination of facet and GPL processes through the relationship

$$f^3(x)=\frac{af^1(x)+bf^2(x)}{\sqrt{a^2+b^2}}, \tag{17}$$

where $f^1(x)$ is the facet process and $f^2(x)$ is the GPL process. This relation ensures that the hybrid process has the same PSD as the GPL and facet processes and allows its scale structure to be adjusted between feature-like and maximally non-feature-like. In this paper, the constants are constrained to $a=\sqrt{3}$ and $b=1$ such that 75% of the energy comes from the facet process.

A realization of the hybrid process is shown in yellow in Fig. 1. It appears qualitatively non-feature-like, just like the GPL realization, in spite of drawing 75% of its energy from the feature-like realization. This example raises two important questions. First, there is the practical issue of the detectability of feature-like signals amidst non-feature-like noise. A geologist studying the hybrid profile of Fig. 1 would be keenly interested in the hidden feature-like signal because it might be readily ascribed to a particular formation mechanism. Yet neither its qualitative appearance nor a second-moment analysis would suggest a departure from a Gaussian, maximally non-feature-like, fractal power law model and therefore the important feature-like signal would probably go undetected. The failure of the qualitative and PSD approaches to scale structure determination in the case of the hybrid surface raises the second issue, which is the need for a quantitative definition of scale structure that would allow a continuum of values. The wavelet transform is proposed to be a useful tool to address both problems since, like scale structure, it is based on a space–scale concept.

The usefulness of the deterministic continuous wavelet transform in detecting scale structure is studied in Fig. 5. The last three realizations of Fig. 1 are reproduced; the background is a color contour plot representing the logarithm of the energy in the continuous wavelet transform as a function of scale (y axis) and space (x axis), using the Daubechies 2 wavelet. The x axis indicates where on the surface the analyzing wavelet is centered, while the y axis corresponds roughly to the length of the analyzing wavelength and applies only to the contour plots of the wavelet transform, not to the plots of the realizations. Both axes are normalized by correlation length.

The dark areas in the case of the facet realization correspond to those areas where the wavelets lie entirely within one of the facets and yield zero energy because the Daubechies 2 wavelet is orthogonal to linear variations. If a wavelet at a given scale is smaller than a facet, all smaller-scale wavelets centered at that position also yield zero energy, showing a dependence across scale.

Neither the GPL nor the hybrid cases show a dependence across scale. Instead, the loci where the wavelets yield zero energy form a complex pattern of fractal bifurcations. Figure 5 demonstrates that the deterministic continuous wavelet transform of realizations does not provide any more ability to detect scale structure differences than qualitative perusal of the realizations themselves.

The statistics of wavelet transform coefficients are more helpful. Since the wavelet transform is a linear operator, the second moment statistics of wavelet coefficients depend only on the second moment statistics of the surfaces. Therefore, in using wavelet coefficient statistics, the only way to detect scale structure differences between processes with identical correlation function is to consider moments greater than the second.

Among the many combinations of higher moments from the multidimensional joint pdf of all coefficients across scale and space, only the full one-dimensional pdf of the coefficients at a single scale is considered here. Figure 6 displays histograms of the wavelet coefficients at a scale of one-tenth of a correlation length using 50 realizations for each of the surface types in Fig. 5. The histograms are normalized to unit variance. The blue curves in each plot represent a zero-mean, unit variance Gaussian pdf showing that the GPL

model yields Gaussian wavelet coefficients, the facet case yields highly non-Gaussian coefficients, and the hybrid model lies between the two extremes.

The one-dimensional pdf of coefficients at a single scale has successfully distinguished all three surface types in this example in which all models have the same PSD. A word of caution is warranted lest it appear that a good quantitative definition of scale structure has been found. An alternate definition of Gaussianity for a stochastic process $f(x)$ is that the scalar

$$a=\int_{-\infty}^{\infty} f(x)w(x)\,dx \tag{18}$$

is Gaussian for all L^2-integrable choices of $w(x)$.[8] Since the wavelet transform is of the form in Eq. (18) with $w(x)$ equal to the analyzing wavelet and a equal to the corresponding wavelet coefficient, it follows that the one-dimensional coefficient pdf at a single scale is, in fact, a measure of Gaussianity rather than scale structure. All Gaussian processes lead to Gaussian one-dimensional coefficient pdf's irrespectively of their scale structure. Thus, while being feature-like, the GG process would yield a normalized histogram identical to the maximally non-feature-like GPL process. Nevertheless, the current example shows that Gaussianity can serve as an indirect measure of scale structure when the PSD is known.

To fully incorporate the scale-space nature of scale structure, a statistical metric should at least provide information regarding the dependence of co-located coefficients across scale. A satisfactory analytical definition of scale structure is not yet available, but wavelets appear to be a promising approach.

II. THEORY

A. Integral equation solution

The Helmholtz–Kirchhoff integral formula gives the total acoustic pressure in a volume V as an integral of source terms distributed over a smooth closed surface bounding V. This formula can be adapted for one-dimensional, rigid, finite-length surfaces S, yielding[1,10]

$$p(\mathbf{r})=p_{\mathrm{inc}}(\mathbf{r})+\frac{1}{4i}\int_S p(\mathbf{r}')\,\frac{\partial H_0^{(1)}(k|\mathbf{r}-\mathbf{r}'|)}{\partial \mathbf{n}'}\,dS'. \tag{19}$$

Equation (19) is valid for any $\mathbf{r}$ not on S. In the limit where $\mathbf{r}\to\mathbf{r}''\in S$, Eq. (19) becomes

$$\frac{1}{2}p(\mathbf{r}'')=p_{\mathrm{inc}}(\mathbf{r}'')+\frac{1}{4i}\int_S p(\mathbf{r}')\,\frac{\partial H_0^{(1)}(k|\mathbf{r}''-\mathbf{r}'|)}{\partial \mathbf{n}'}\,dS'. \tag{20}$$

Equation (20) is a Fredholm integral equation of the second kind for the *exact* total pressure $p(\mathbf{r}'')$ on the surface, including all multiple scattering and shadowing effects.[4] For arbitrary S, this equation must be solved numerically by discretization. It can be proven that the discretized version of a Fredholm equation of the second kind converges to a unique and correct result as the discretization interval tends to zero.[11]

The surface is sampled at equal intervals Δx and Eq. (20) becomes a matrix equation:

$$\mathbf{a}=\mathbf{H}\mathbf{b}, \tag{21}$$

where $\mathbf{b}_l=p(\mathbf{r}_l)$ is the unknown vector, $\mathbf{a}_k=p_{\mathrm{inc}}(\mathbf{r}_k)$, and

$$\mathbf{H}_{kl}=\begin{cases}\dfrac{i}{4}\Delta x\,\gamma(\mathbf{r}_l)\dfrac{\partial H_0^{(1)}(k|\mathbf{r}_k-\mathbf{r}'|)}{\partial \mathbf{n}'}\Big|_{\mathbf{r}'=\mathbf{r}_l}, & k\neq l,\\ \dfrac{1}{2}-\dfrac{\Delta x}{4\pi\gamma(\mathbf{r}_k)^2}\dfrac{d^2f(x')}{dx'^2}\Big|_{\mathbf{r}'=\mathbf{r}_k}, & k=l,\end{cases} \tag{22}$$

where $\gamma^2(\mathbf{r}')=1+(df(x')/dx')^2$ and $f(x)$ is the function defining the surface S. The normal derivative of the Hankel function in Eq. (22) is

$$\frac{\partial H_0^{(1)}(k|\mathbf{r}''-\mathbf{r}'|)}{\partial \mathbf{n}'}=\frac{kH_1^{(1)}(k|\mathbf{r}''-\mathbf{r}'|)}{\gamma(x')|\mathbf{r}''-\mathbf{r}'|}[(f(x')-f(x''))-f'(x')(x'-x'')]. \tag{23}$$

For the diagonal elements of $\mathbf{H}$ where $\mathbf{r}''=\mathbf{r}'$, $f(x'')$ is expanded in a Taylor series about x' as suggested in Ref. 4, and the limit is taken as $\mathbf{r}''\to\mathbf{r}'$, yielding

$$\lim_{\mathbf{r}''\to\mathbf{r}'}\frac{\partial H_0^{(1)}(k|\mathbf{r}''-\mathbf{r}'|)}{\partial \mathbf{n}'}=\frac{if''(x')}{\pi\gamma^3(x')}. \tag{24}$$

The scattered pressure, defined as $p_s(\mathbf{r})=p(\mathbf{r})-p_{\mathrm{inc}}(\mathbf{r})$, is computed at any location $\mathbf{r}$ in the fluid from the discretized version of Eq. (19):

$$p_s(\mathbf{r})=\sum_k \frac{\mathbf{b}_k\Delta x\,\gamma(\mathbf{r}_k)}{4i}\,\frac{\partial H_0^{(1)}(k|\mathbf{r}-\mathbf{r}'|)}{\partial \mathbf{n}'}\Big|_{\mathbf{r}'=\mathbf{r}_k}. \tag{25}$$

B. Narrowband fields

A tapered plane wave incident field identical to that in Ref. 4 is used:

$$p_{\mathrm{inc}}(\mathbf{r})=\exp\{i\mathbf{k}_{\mathrm{inc}}\cdot\mathbf{r}[1+w(\mathbf{r})]-(x-z\cot\theta)^2/g^2\}, \tag{26}$$

where $w(\mathbf{r})=[2(x-z\cot\theta)^2/g^2-1]/(kg\sin\theta)^2$. The incident grazing angle is θ and the incident wavenumber is $\mathbf{k}_{\mathrm{inc}}=k(\cos(\theta+\pi),\sin(\theta+\pi))$ as shown in Fig. 7. The incident field has Gaussian amplitude taper perpendicular to the incidence direction. The taper suppresses the edge effects that would otherwise arise from the use of a finite-length surface instead of a closed surface in the Helmholtz–Kirchhoff integral. A consequence of the chosen taper is that the incident field is only an approximate solution to the Helmholtz equation. The particular form used satisfies the Helmholtz equation to order $1/(kg\sin\theta)^2\ll 1$.[4] The taper parameter g is set to $L/4$ for all surfaces considered here, where L is the horizontal extent of the surface.

A useful quantity in distinguishing the angular dependence of narrowband scatter from various one-dimensional surfaces is the scattering cross section, defined as the mean-square scattered pressure evaluated in the far field normal-

ized by the incident energy and a cylindrical spreading factor. The far-field scattering strength SS is ten times the logarithm of the scattering cross section. For the incident field in Eq. (26),

$$\mathrm{SS}(\theta,\theta_s) = \lim_{|\mathbf{r}|\to\infty} 10\log\frac{|\mathbf{r}|\mathcal{E}[|p_s(\mathbf{r})|^2]}{\sqrt{\pi/2}g[1-0.5(1+2\cot^2\theta)/(kg\sin\theta)^2]}, \tag{27}$$

where $p_s(\mathbf{r})$ is found from Eq. (25) numerically. Since the sample surfaces scale with wavelength, the scattering strength is independent of frequency and sound speed.

C. Wideband fields

Central to this paper is an exploration of the time-domain statistics of scattered real bandlimited pulses with spectrum

$$G(\omega)=\begin{cases} G_b(\omega), & 1-\Gamma/2\leqslant\omega/\omega_c<1+\Gamma/2,\\ G_b^*(-\omega), & -1-\Gamma/2\leqslant\omega/\omega_c<-1+\Gamma/2,\\ 0, & \text{otherwise,}\end{cases} \tag{28}$$

where Γ is the proportional bandwidth, defined as the ratio of bandwidth to center frequency ω_c.

The time-domain response of each surface realization is obtained from the inverse Fourier transform of $p_s(\mathbf{r},\omega)$,

$$p_s(\mathbf{r},t)=\frac{1}{2\pi}\int_{-\infty}^{\infty} p_s(\mathbf{r},\omega)e^{-i\omega t}\,d\omega, \tag{29}$$

where $p_s(\mathbf{r},\omega)$ is obtained by solving Eqs. (21) and (25) at each frequency using the wideband incident field

$$p_{\mathrm{inc}}(\mathbf{r},\omega)=p_{\mathrm{inc}}(\mathbf{r})G(\omega). \tag{30}$$

As a time-domain analog to the scattering strength, the far-field zero-mean log-envelope $E(t,\theta,\theta_s)$ is defined as

$$E(t,\theta,\theta_s)=\lim_{|\mathbf{r}|\to\infty}(10\log|p_s(\mathbf{r},t)|^2 - \mathcal{E}[10\log|p_s(\mathbf{r},t)|^2]). \tag{31}$$

Since $\theta_s=\theta=45°$ for all cases considered here, the notation for the log-envelope is simplified to $E(t)$.

The statistics of time-domain backscatter for each surface type are characterized using the pdf of E, $p_E(E)$. Note

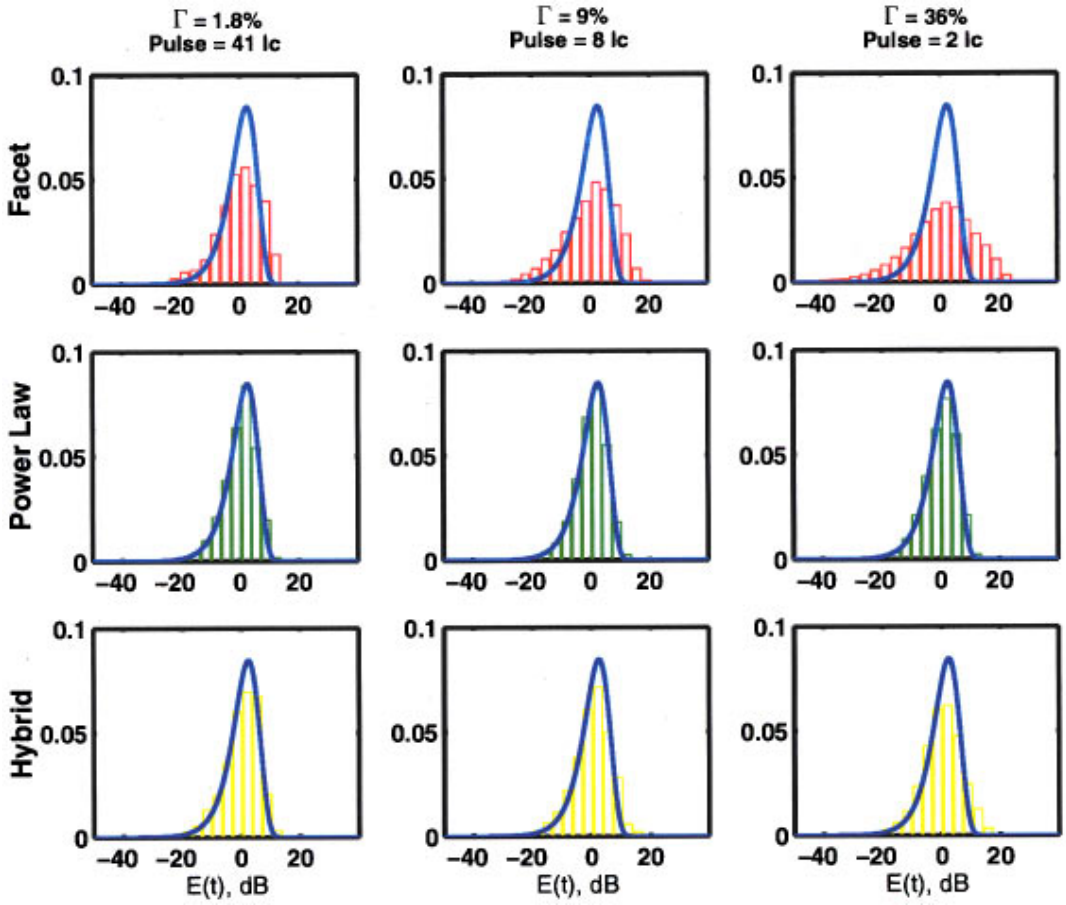


FIG. 9. Estimates of the pdf's of the log-envelope $E(t)$ for $(\sigma,l_c)=(0.3,4.0)\lambda$. The log-transformed Rayleigh distribution is shown in dark blue, and the pulse resolution is shown in correlation lengths for each bandwidth. $E(t)$ is nearly Rayleigh at all proportional bandwidths Γ for the GPL surface because the multi-scale roughness presents many scatterers within each pulse width. The facet surface deviates increasingly from Rayleigh as the bandwidth is increased and the pulse is better able to resolve individual back-reflecting facets. The presence of 25% non-feature-like energy in the hybrid surface leads to a reduction in the deviations from Rayleigh. Nevertheless, at all three values of bandwidth the underlying feature-like ''signal'' is felt by the acoustics in spite of being invisible to the eye and undetectable through spectral methods.

that $p_E(E)$ is independent of ω_c and sound speed c since a change in either parameter simply temporally scales each backscattered time series.

A limiting case for $p_E(E)$ occurs when the backscatter $p_s(t)$ is the sum over a large number of independent, identically distributed variables (scatterers). By the central limit theorem, the quantity $r(t)=|p_s(t)|$ then has a Rayleigh density:

$$p_r(R)=\frac{R}{\sigma_r^2}e^{-R^2/2\sigma_r^2}u(R), \tag{32}$$

where $u(R)$ is the unit step function. The probability density of the resulting log-envelope, $E_R(t)$, is

$$p_{E_R}(E_R)=\frac{\ln 10}{20\sigma_r^2}10^{E_R/10}\exp(-10^{E_R/10}/2\sigma_r^2). \tag{33}$$

This density will be referred to as *log-transformed Rayleigh*. The fact that E_R is zero-mean by definition constrains the value of σ_r^2 as follows:

$$0=\mathcal{E}[E_R]=\int_0^\infty 20\log R\,\frac{R}{\sigma_r^2}e^{-R^2/2\sigma_r^2}\,dR$$

$$=10\log e(\ln 2\sigma_r^2-C), \tag{34}$$

implying that $\sigma_r^2=e^C/2\simeq 0.89$, where C is Euler's constant (0.577 215...).[12]

In this paper, $p_E(E)$ is estimated numerically for each surface type by forming a histogram of 50 backscattered log-envelopes. The individual time series are truncated to eliminate sidelobe leakage and are concatenated. Regarding the accuracy of the pdf's estimated using this Monte Carlo approach, note that all comparisons involving estimated pdf's are warranted by chi-square tests using the number of independent observations as the number of degrees of freedom. For each surface, this number is roughly equivalent to the number of pulse widths fitting within the length of the surface, $L/\Delta x_p$. By concatenating the 50 realizations used in each experiment, one gets a minimum of 50 independent observations of $E(t)$, which is achieved when the pulse's spatial extent on the surface is greater than or equal to L.

III. RESULTS AND DISCUSSION

In this section, the four models of Sec. I are used in two numerical experiments. In the first experiment (Sec. III A), the narrowband bistatic scattering strength and the log-envelope histograms of backscattered wideband pulses are compared for facet, GPL, and hybrid surfaces. In the second experiment (Sec. III B) they are compared for GG, facet, and GPL surfaces.

In each simulation the exact scattered field is computed using the numerical method of Sec. II. Ensemble averages are obtained using 50 sample surfaces each of which is rigid, one-dimensional, and 80 wavelengths long. The incident grazing angle is 45 degrees.

As discussed in Appendix B, the GPL, hybrid, and facet surfaces are low-pass filtered to remove slope discontinuities for compatibility with the integral equation. Appendix B shows that the acoustic properties of the surfaces are not altered by the filtering operation.

The pulse used in the wideband simulations has a spectrum $G_b(\omega)$ [see Eq. (28)] equal to a Chebyshev window of order 50 and bandwidth $\Gamma\omega_c$. This pulse has constant sidelobe level equal to -50 dB and its time-bandwidth product is $\Delta t\Delta\omega/(2\pi)\simeq 4.2$. From this, one can deduce the spatial extent of the pulse in wavelengths:

$$\frac{\Delta x_p}{\lambda}=\frac{c\Delta t}{2\lambda\cos\theta}\simeq\frac{2.1}{\Gamma\cos\theta}. \tag{35}$$

Three values of proportional bandwidth Γ are used: 1.8%, 9%, and 36%.

A. Scatter from facet, GPL, and hybrid surfaces

In this experiment all three surface models have the same PSD but, as discussed in Sec. I, their scale structure differs, enabling the study of the sufficiency of the second moment in predicting the ensemble scattering properties of surfaces, and the additional role played by scale structure beyond the second moment. The rms height and correlation length are fixed at $(\sigma,l_c)=(0.3,4.0)\lambda$. In the wideband case, the normalization wavelength λ is the wavelength at the *center frequency*.

1. Narrowband bistatic scattering strength

Shown in Fig. 8 are the bistatic scattering strength curves for the three power-law surfaces, and also the curve for the GG surface which is discussed in Sec. III B. The facet, GPL, and hybrid curves are distinct from one another, proving that the second moment is not sufficient to predict the average scattering strength of a stochastic surface. Figure 8 provides a quantitative look at the kinds of gains in predictions of scattering strength which could be realized by accounting for scale structure. The facet surface exhibits a different functional dependence than the GPL surface, manifesting a 3–5-dB increase about the specular direction and a 3–6-dB decrease in the back quadrant, while the hybrid curve lies roughly half-way between the other two curves at all angles. Overall, the maximum difference is about 6 dB between the GPL and facet models at near-normal grazing angles.

The fact that the hybrid curve is distinct from the GPL curve demonstrates that the acoustics are sensitive to changes in scale structure which are undetectable either through spectral methods or by qualitative analysis of sample realizations. It provides physical justification for the idea of scale structure as a continuous parameter which should be characterized quantitatively.

The next section shows that scale structure plays an even more dramatic role in the time-domain.

2. Wideband time-domain backscatter statistics

Having established that scale structure plays a role in mean scatter energy, the focus now turns to the temporal fluctuations of backscattered log-envelopes about their means, which are characterized using $p_E(E)$.

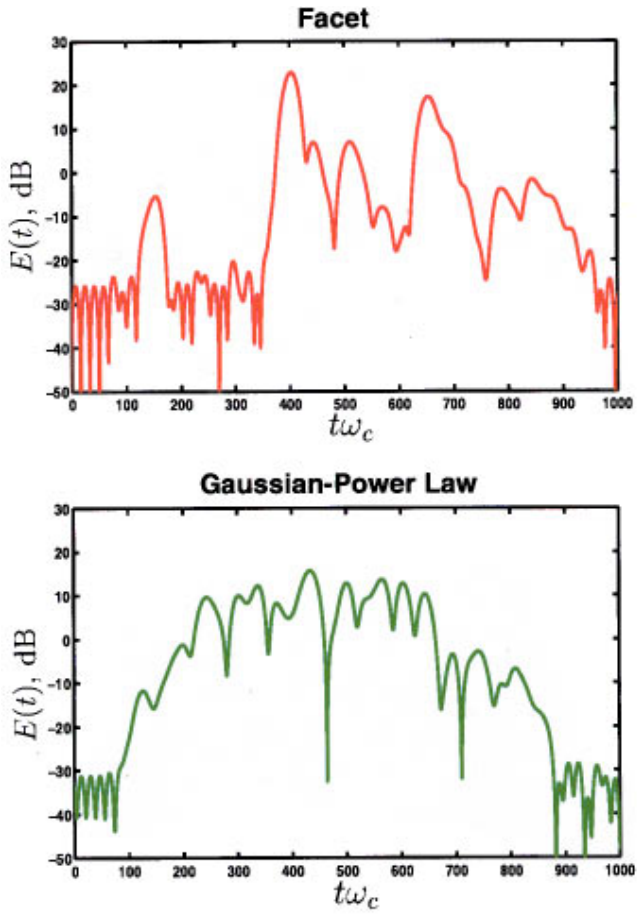


FIG. 10. Sample realizations of the log-envelope for $(\sigma,l_c)=(0.3,4.0)\lambda$ and a proportional bandwidth of 36%. The facet surface leads to log-envelopes with strong peaks separated by quiescence. The GPL surface does not give rise to well-separated events.

The results at bandwidths of 1.8%, 9%, and 36% are presented in the form of histograms in Fig. 9. The solid line overlaid on each histogram is the log-transformed Rayleigh distribution given by Eq. (33). The spatial extent of the incident pulse is indicated in units of correlation length at the top of each column; the sample paths of Fig. 1 may be helpful to gain an intuitive feel of how much of the surface is included by the pulse at each bandwidth.

Scatter from the facet surface in the first row deviates from Rayleigh at all bandwidths and the deviations increase with bandwidth. Looking at the second row, the envelope statistics of the GPL surface are perfectly Rayleigh at proportional bandwidths of 1.8% and 9% and are nearly so at 36%. In the third row, the fluctuations of backscatter for the hybrid surface deviate from the log-transformed Rayleigh curve, but not as much as the facet case.

To understand the dramatic effect of scale structure on the fluctuations, it is insightful to look at the two sample log-envelopes for the 36% case shown in Fig. 10. Scatter from the facet surface shows short high-amplitude bursts separated by quiescent periods. As a consequence of these *event-like* or *target-like* fluctuations, the pdf's for the facet surface show enhanced low- and high-level tails and a depressed central region, reflecting the tendency of the log-envelope to spend more time at extreme levels and less time

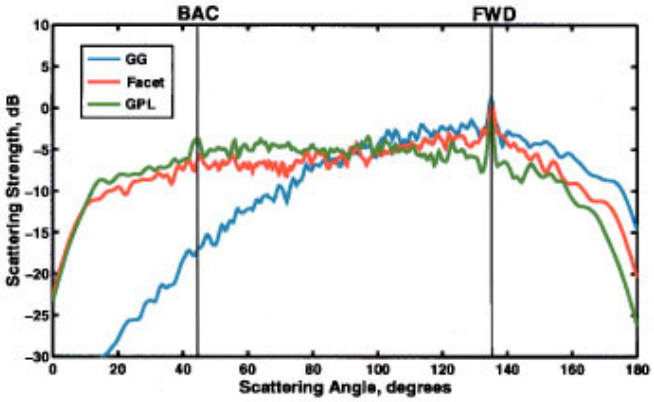


FIG. 11. Monte Carlo estimates of scattering strength for $(\sigma,l_c)=(0.2,1.0)\lambda$. Once again, the power-law PSD of the facet and GPL surfaces leads to enhanced backscatter compared to the GG surface and the feature-like GG and facet models show an enhanced forward scatter lobe compared to the GPL model. In all cases, energy conservation is satisfied to within 1%.

at mean levels compared with a Rayleigh process. This behavior suggests that the spatial extent of the pulse, which from Eq. (35) is $\Delta x_p \simeq 2l_c$ at a bandwidth of 36%, is resolving individual targets on the surface.

This theory can be verified from the following argument. The average facet size is well over a wavelength, so the facets are likely to behave as reflectors. The average number of facets per pulse, given by $\Delta x_p \lambda_r$, is about four as calculated from the formula

$$\Delta x_p \lambda_r = \frac{10}{3\sqrt{3}} \frac{\Delta x_p}{l_c}, \tag{36}$$

where Eq. (12) was used. At only four facets per pulse it is unlikely that more than one retroreflecting facet will be insonified at a time, hence the event-like backscatter is seen to be caused by *occasional* glints from back-facing facets. This effect is diminished as the bandwidth is decreased because the probability of many glints occurring at the same time is increased.

In sharp contrast to the facet case, the sample envelope from the GPL surface in Fig. 10 does not give rise to clearly distinguishable events. From the 36% bandwidth column in Fig. 9, the statistics of the envelope are almost perfectly Rayleigh, suggesting that the scattered pressure is the sum over a sufficiently large number of contributions for the central limit theorem to apply. Physically, these contributions arise in the form of scatterers lying within the pulse's spatial resolution. Each part of the GPL surface, because of the maximally non-feature-like scale structure, presents scatterers of all sizes scrambled spatially, which hamper the ability of presenting locally coherent scattering separated by relative quiescence.

This analysis of the fluctuations from the facet and GPL models demonstrates that, given a common second moment, surface models with different scale structure can scatter through fundamentally different physical mechanisms. The time-domain results for the hybrid surface emphasize even more strongly than the scattering strength results that acous-

TABLE I. Value of the rms slope angle γ for the surfaces used in this study. As described in Appendix B, the facet, GPL, and hybrid surfaces have been low-pass filtered using a Kaiser–Bessel window of order 5 with $k_c = 4k$.

$(\sigma, l_c)/\lambda$	Gaussian–Gaussian (degrees)	Gaussian–power-law, facet and hybrid (degrees)
(0.2, 1.0)	15.8	28.1
(0.3, 4.0)	6.1	18.4

tics are sensitive to changes in scale structure not easily detectable by conventional means, i.e., using the human eye or second moment analyses.

B. Scatter from facet, GG, and GPL surfaces

This second set of numerical experiments studies whether a qualitative determination of scale structure provides at least some predictive ability without knowledge of the PSD. To this end, the narrowband bistatic scattering strength and the log-envelope histograms of backscattered wideband pulses are compared for the facet and GG surfaces, which have feature-like scale structure but quite different PSDs. The calculations are also performed for the GPL surface, therefore this investigation includes all three poles of the surface property triangle in Fig. 2. Along with the results from the first experiment, this triad of surfaces helps separate scattering behavior governed by scale structure from that governed by the PSD.

1. Narrowband scattering strength

The expected value of the bistatic scattering strength is estimated for two cases: $(0.3, 4.0)\lambda$ in Fig. 8 and $(\sigma, l_c) = (0.2, 1.0)\lambda$ in Fig. 11. Two observations are true regarding both figures: (i) the facet and GPL surfaces, sharing a common power-law PSD, lead to greatly enhanced scatter in the back quadrant compared to the GG surface; (ii) the facet and GG surfaces, sharing a feature-like scale structure, show enhanced scatter within ± 20 degrees of the specular direction compared to the GPL surface.

To analyze these observations, first note that, for the GG surface, the validity of the Kirchhoff approximation at both values of (σ, l_c) (Ref. 4) suggests that the total scattering strength in any direction arises from locally plane segments in reflection. The most likely slopes are those near zero, hence there is a bulge of incoherent energy about the specular direction. Scatter in the back quadrant is weak for the GG surface since steep, back-facing slopes are not common at $(\sigma, l_c) = (0.2, 1.0)\lambda$, the rms slope angle being 15.8 degrees from Table I, and because in the absence of diffraction there is no significant mechanism to inject energy in that quadrant. At $(\sigma, l_c) = (0.3, 4.0)\lambda$, the rms slope angle of 6.1 degrees is even lower, explaining the corresponding decrease of energy in the back quadrant in Fig. 8.

On the other hand, the GPL surface never presents locally plane segments because it contains structure down to subwavelength scales. Each tiny feature is a diffractor which radiates in all directions. The total scatter is the cumulative effect of these diffractors, which explains the quasi-omnidirectional beampattern of the GPL surface where backscatter is no more than 5 dB weaker than incoherent forward scatter.

The facet surface is unique in that it consists of a combination of reflectors (the facets) and diffractors (the vertices). The reflectors, whose slope distribution is peaked about zero, lead to enhanced scatter about the forward direction as in the GG surface. The vertices and the facets with sufficiently high slopes, which are represented by the high-wavenumber tail of the PSD, lead to enhanced scatter in the back quadrant. As a result, the scattering strength of the facet surface resembles that of the GG surface about the specular direction and that of the GPL surface everywhere else.

Based on these interpretations one arrives at the following propositions regarding the roles of scale structure and the PSD:

(1) The ability to produce a forward scatter lobe in a surface whose slope distribution is peaked about zero is determined by the presence of smooth features which must be large enough to act as local reflectors. *Scale structure* is thus a good predictor of the existence or inexistence of a forward scatter lobe.
(2) The ability to produce high scatter in the back quadrant is governed by high-wavenumber roughness, which takes the form of sharp corners and short steep slopes. Thus, the *power spectral density* is an indicator of high energy in backscatter.

An important task in future efforts will be to see how these propositions hold at other values of (σ, l_c), θ, and for other types of random surfaces.

Given that two of the axes in the feature triangle of Fig. 2, namely the PSD and scale structure, are observed to be useful characterizers of roughness in acoustic scattering, it is natural to ask whether the third axis, Gaussianity, is also useful. The answer appears to be that it is not. While the absence of commonalities in the behavior of the two Gaussian surface models (GG and GPL) does not necessarily rule out this proposition, a physical justification for the usefulness of Gaussianity as an independent parameter is improbable.

2. Wideband time-domain backscatter statistics

In the time-domain results of Sec. III A, a tendency for feature-like scale structure to enhance event-like fluctuation statistics was observed. Further evidence is provided in Fig. 12, where time-domain backscatter statistics for the GG, facet, and GPL surfaces are presented for $(\sigma, l_c) = (0.2, 1.0)\lambda$. The two feature-like interfaces, in spite of the vast differences in their PSDs, are the only ones to exhibit non-Rayleigh statistics.

The perfectly Rayleigh results for the GPL surface are easily explained as before by the multi-scale feature distribution, but the behavior of the GG and Facet surfaces merit further discussion. The first observation is that feature-like roughness is *not sufficient* to ensure event-like scatter since the facet surface is perfectly Rayleigh at $\Gamma = 1.8\%$. In fact, at low enough bandwidths, the central limit theorem ensures that all rough surfaces lead to Rayleigh statistics because arbitrarily many scatterers are included in the response. In-

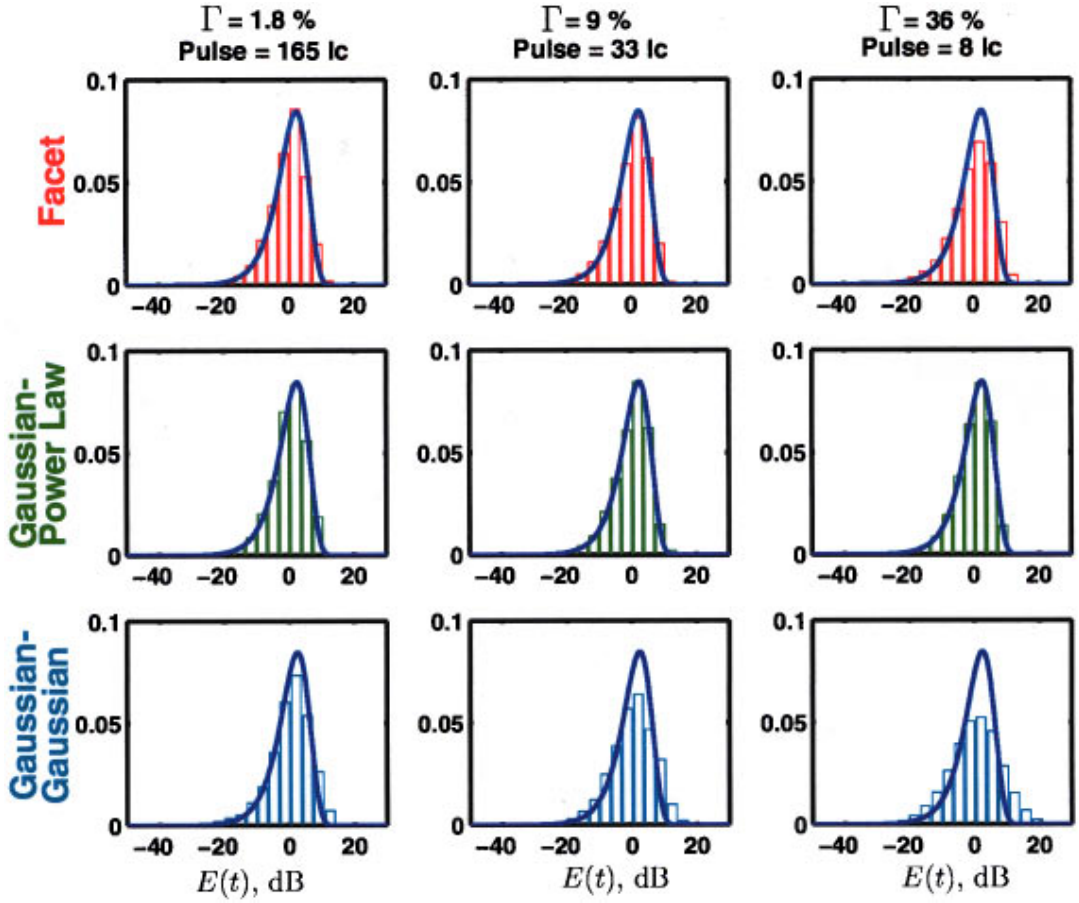


FIG. 12. Estimates of the pdf's of the log-envelope for $(\sigma, l_c) = (0.2, 1.0)\lambda$. The log-transformed Rayleigh distribution is shown in dark blue, and the pulse resolution is shown in correlation lengths for each bandwidth. The feature-like GG and facet models both lead to non-Rayleigh statistics while the GPL model is perfectly Rayleigh at all bandwidths. The GG surface is event-like at lower proportional bandwidths than the Facet surface because of its lower rms slope angle γ and the larger size of its features.

terestingly, feature-like roughness is *not necessary* either for giving rise to event-like statistics. Indeed, the $\Gamma = 36\%$ histogram for the maximally non-feature-like GPL surface in Fig. 9 shows a depressed central peak which, although small, is warranted to be statistically significant by a chi-square test. Also, simulations for the GPL surface at $(\sigma, l_c) = (2.0, 8.0)\lambda$ in Ref. 1 show strong event-like behavior. (As expected, scatter from the corresponding facet surface is even more event-like.)

While feature-like scale structure is neither necessary nor sufficient for event-like scatter, it clearly plays an important role in enhancing it compared with non-feature-like surfaces. While useful, this observation is not quantitative enough to explain differences in scattering statistics between feature-like surfaces. It does not explain (i) why the onset of event-like scatter occurs at a lower bandwidth in the GG surface than in the facet surface in Fig. 12 or (ii) why the statistics for the facet surface at $\Gamma = 9\%$ in Fig. 9 deviate much more from Rayleigh than at $\Gamma = 36\%$ in Fig. 12, despite the fact that there are 16 facets per pulse in each case. Both questions can be answered with a more detailed analysis of the scale structure of each surface.

To answer question (i), first note that the GG surface, with an rms slope angle γ of 15.8 degrees, presents back-facing slopes much less frequently than the facet surface, for which $\gamma = 28.1$ degrees. (In the facet surface, the rms slope angle is equivalent to the rms slope angle of the facets themselves.) Thus in the facet surface the envelope is not as likely to experience quiescence between glints and is more likely to include more than one glint at a time, improving the match of the statistics with a Rayleigh process. A second reason for the higher deviations from Rayleigh in the GG surface is that the average width of a feature, being equal to l_c, is about twice that of an average facet $(1/\lambda_r \simeq 0.52 l_c)$. As a result, there are half as many features per pulse and thus better defined glints in the GG surface.

Similar arguments help answer question (ii). The rms slope angle in Fig. 9, being 18.4 degrees, leads to rarer glints than Fig. 12 where $\gamma = 28.1$ degrees. Incidentally, the relative rarity of the glints at $(\sigma, l_c) = (0.3, 4.0)\lambda$ reduces the mean backscatter levels by about 7 dB compared with $(\sigma, l_c) = (0.2, 1.0)\lambda$. Similarly, the even rarer glints of the GG surface help explain why its backscatter is lower than the facet surface's.

These analyses demonstrate both the usefulness of scale structure and its limits as a qualitative tool that supplements the PSD. An analytical definition of scale structure, either based on wavelets as suggested in Sec. I D or on some other method, could remove these limitations. Just as the PSD not

only supplements but *embodies* the more basic descriptors of variance and correlation length, the ultimate goal for a quantitative definition of scale structure is for it to both supplement and embody the PSD. It should eventually permit expressions for the acoustic field to be analytically linked with those for scale structure and be able to predict phenomena associated with the organization of features as reliably as those associated with the PSD.

IV. CONCLUSION

Scale structure is a useful new concept that describes the roughness of interfaces according to the spatial arrangement of their component features at various scales. When used as a qualitative tool, scale structure can provide valuable insights into the qualitative scattering behavior of rough surfaces. For example, in this paper it was found that the surfaces with feature-like scale structure exhibited a forward scatter lobe and enhanced event-like backscatter compared with the maximally non-feature-like surface.

The power spectral density (PSD) performed well as a qualitative predictor of scattering strength in the back quadrant irrespectively of scale structure. However, the distinctness of the time-domain statistics and the forward quadrant scattering strength in surfaces with identical PSD but different scale structure highlighted the insufficiency of second moment characterizations in capturing all the acoustically important information about roughness. Scale structure is an effective concept because it is rooted in the physical observation that *features*, not wavenumber components or statistical moments, are the most direct cause of scattering.

The qualitative approach to determining scale structure has its limitations, highlighting the need for a quantitative definition. It was found that while two surface models classified as ''feature-like'' showed enhanced event-like scatter, the description was too imprecise to predict the significant differences in scattering between them. In another example, a hybrid model was presented whose realizations were found to be visually indistinguishable from the realizations of a non-feature-like model despite drawing 75% of their energy from a feature-like model. The qualitative appearance could not be relied on since the acoustics yielded bistatic scattering strengths and time-domain statistics which were distinct. This result also demonstrated the sensitivity of sound to a *continuum* of variations in scale structure. A method for quantifying scale structure was therefore deemed necessary not only for proper acoustical prediction but also as a means of detecting feature-like signals buried in non-feature-like noise. Wavelets were shown to be a promising approach.

The results in this paper have made a strong case that it is worthwhile for the acoustician to examine the scale structure of rough surfaces. The author is hopeful that others will be inspired to explore this area and break new ground. Many topics for further study are apparent, including (i) the role of scale structure at other values of variance and correlation length; (ii) its role in two-dimensional surface scattering and volumetric scattering; (iii) its effect on scattering from elastic media; (iv) the possibility of modifying current scattering theories to include scale structure effects; and finally (v) the quantitative definition of scale structure and the analytical connection of expressions for scale structure with expressions for scattering strength and time-domain fluctuation statistics.

ACKNOWLEDGMENTS

The author is grateful to Professor Arthur Baggeroer for guidance and financial support, Professor Henrik Schmidt for technical advice, Professor Ira Dyer for inspiration, and Professor Clarence Clay for illuminating comments. Nancy Lum, Dr. Joseph Bondaryk, and Dr. Pierre Elisseeff provided insightful reviews, as did an anonymous reviewer. This work was supported by the Acoustical Reverberation Special Research Program of the Office of Naval Research.

APPENDIX A: DEFINITION OF CORRELATION LENGTH

The definition of correlation length used in this paper is

$$l_c = \sqrt{\frac{2\mu_2}{\mu_0}}, \tag{A1}$$

where

$$\mu_0 = 2\int_0^\infty R_{hh}(\chi)\,d\chi \tag{A2}$$

and

$$\mu_2 = 2\int_0^\infty \chi^2 R_{hh}(\chi)\,d\chi. \tag{A3}$$

APPENDIX B: ISSUES IN NUMERICAL IMPLEMENTATION

1. Surface filtering and convergence

Recall from Sec. II that the integral equation formulation requires each scattering surface to be smooth, i.e., that its first derivative be continuous everywhere. The GPL, hybrid, and facet surfaces all violate this requirement and must be appropriately modified for use with the numerical method. A surface $f(x)$ with any number of slope discontinuities can be made smooth by low-pass filtering,

$$f_d(x) = \int_{-\infty}^{\infty} f(x')\,d(x-x')\,dx', \tag{B1}$$

where $d(x)$ is the impulse response of a low-pass filter and $f_d(x)$ is the filtered surface. If the passband in the low-pass filter extends to a sufficiently high corner wavenumber k_c, the smoothed surfaces are acoustically equivalent to their unsmoothed versions because the features being smoothed are too small compared to the wavelength to be physically relevant. In these situations the scatter is said to have ''converged in k_c.'' The scatter from a surface with slope discontinuities can be computed as the limit of scatter from a sequence of smooth, low-pass filtered surfaces with increasing corner wavenumber.

It is customary to design the low-pass filter $D(k)$ in the wavenumber domain, aiming for a sharp drop in response beyond the corner wavenumber. Such filters lead to impulse responses with large ripples that spatially distort the scatter-

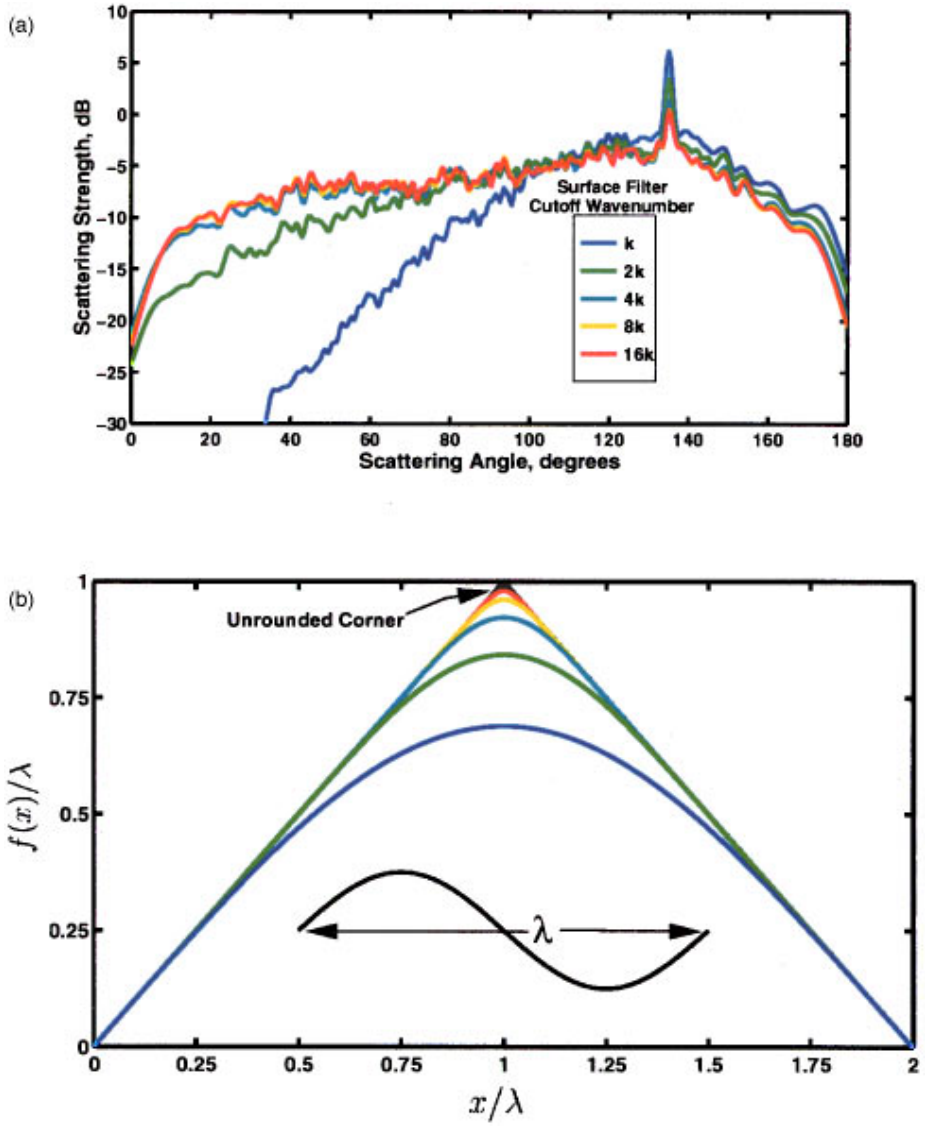


FIG. B1. (a) Monte Carlo estimates of average scattering strength of low-pass-filtered surfaces with $(\sigma,l_c)=(0.2,1.0)\lambda$ and increasing values of the corner wavenumber k_c. Convergence occurs at $k_c=8k$, meaning that for $k_c\geqslant 8k$ the filtered surfaces are acoustically equivalent to their unfiltered versions. (b) Size of a rounded corner relative to the wavelength for increasing values of k_c.

ing surfaces. To avoid spatial domain ripples, the filters are defined by setting their impulse responses $d(x)$ to conventional signal processing windows. In particular, Kaiser–Bessel windows of order 5 are used in which the corner wavenumber k_c is defined as the location of the first null of $D(k)$. The spatial extent L_{KB} of the Kaiser–Bessel window of order 5 is related to k_c through the approximate formula $L_{\mathrm{KB}}\simeq 3.8\pi/k_c$.

To visualize the effect of corner wavenumber on scattering strength, Fig. B1(a) depicts bistatic scattering strength curves for the facet model for $(\sigma,l_c)=(0.2,1.0)\lambda$ at increasing values of k_c, and Fig. B1(b) shows the appearance of a corner after being filtered at the different values of k_c.

Convergence is also an issue with regards to sampling; the filtered surfaces $f_d(x)$ must be sampled finely enough such that the error on the computed scatter is acceptably small. The scatter is said to have ''converged in sampling'' when a doubling in resolution leads to changes in scattering strength which are smaller than or equal to 0.2 dB over all angles of scatter for a given angle of incidence. The sampling rate required for convergence in sampling increases with k_c.

At the values of (σ,l_c) in this paper, which are (0.2, 1.0)λ and (0.3, 4.0)λ, convergence in k_c for the GPL, facet, and hybrid surfaces occurs at $k_c=8k$, where k is the acoustic wavenumber. This value of k_c corresponds to an averaging width of $L_{\mathrm{KB}}\simeq\lambda/4$. Convergence in sampling at this value of k_c occurs at a spacing of $\Delta x=\lambda/16$, requiring $M=1300$ samples for surface lengths of $L=80\lambda$.

The calculations of narrowband bistatic scatter for the GPL, hybrid, and facet models are performed with $k_c=8k$ and $\Delta x=\lambda/16$; the results presented have therefore fully converged. For the wideband time-domain simulations, the computational cost of solving at a sufficiently large number of frequencies to avoid aliasing restricts the investigation to $M\leqslant 800$. The highest corner wavenumber for which 800 samples are sufficient to attain convergence in sampling for the GPL, hybrid, and facet models is $k_c=4k$. Since convergence in k_c occurs at $8k$, the computed backscatter has not quite converged to that of unsmoothed surfaces. The rms difference between the backscattering strengths at $k_c=4k$ and $8k$ is on the order of 1 dB. This error is deemed acceptably low to avoid significant distortions in the backscattered envelopes. Given computational hardware many times faster,

the ideal value of $M=1300$ would be used, which would permit $k_c=8k$.

While the smoothness of the GG surface does not require any low-pass filtering, GG sample surfaces are filtered in the same way as the others for consistency. Convergence in sampling occurs at $\Delta x=\lambda/5$ $(M=400)$ for both values of (σ,l_c). Both the narrow and wideband results presented use this sampling interval and therefore have fully converged.

2. rms slope

A useful quantity in the discussion of the results is the angle of the rms slope. For the GG surface this quantity is $s=\sqrt{2}\sigma/l_c$. For the facet, GPL, and hybrid surfaces, s is undefined. The $1/k^3$ decay of the height spectrum causes a $1/k$ decay of the slope spectrum, leading to a divergent integral. The low-pass filtering necessary for implementation of the numerical method eliminates this logarithmic singularity such that the rms slope exists and is given by

$$s=(\mathcal{E}[(df(x)/dx)^2])^{1/2}$$
$$=\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}k^2S_{ff}(k)|D(k)|^2dk\right)^{1/2}. \quad \text{(B2)}$$

For a corner wavenumber of $4k$, Eq. (41) was solved numerically and the corresponding values of $\gamma=\arctan(s)$ are shown in Table I for all the surfaces used in this study.

Note that while s increases without bound as the corner wavenumber tends to infinity, the existence of a corner wavenumber beyond which all surfaces are acoustically identical suggests a physical definition of the rms slope which can be used for nonintegrable slope spectra.

[1] V. Lupien, ''The importance of scale structure in scattering from random, rough surfaces,'' Ph.D. thesis, Massachusetts Institute of Technology, 1998.
[2] Lord Rayleigh, *The Theory of Sound* (Dover, New York, 1945).
[3] W. A. Kinney and C. S. Clay, ''Insufficiency of surface spatial power spectrum for estimating scattering strength and coherence: Numerical studies,'' J. Acoust. Soc. Am. **78**, 1777–1784 (1985).
[4] E. I. Thorsos, ''The validity of the Kirchhoff approximation for rough surface scattering using a Gaussian roughness spectrum,'' J. Acoust. Soc. Am. **83**, 78–92 (1988).
[5] E. I. Thorsos and D. R. Jackson, ''The validity of the perturbation approximation for rough surface scattering using a Gaussian roughness spectrum,'' J. Acoust. Soc. Am. **86**, 261–277 (1989).
[6] E. I. Thorsos, ''Acoustic scattering from a Pierson-Moskowitz sea surface,'' J. Acoust. Soc. Am. **88**, 335–349 (1990).
[7] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. (McGraw-Hill, New York, 1991).
[8] A. Willsky and G. Wornell, ''Course Notes for 6.432: Stochastic Processes Detection and Estimation,'' Massachusetts Institute of Technology, 1995.
[9] G. W. Wornell, *Signal Processing with Fractals* (Prentice–Hall, Upper Saddle River, NJ, 1996).
[10] J. A. Ogilvy, *Theory of Wave Scattering from Random Rough Surface* (IOP, Philadelphia, 1991).
[11] R. L. Holford, ''Scattering of sound waves at a periodic, pressure-release surface: An exact solution,'' J. Acoust. Soc. Am. **70**, 1116–1128 (1981).
[12] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 5th ed. (Academic New York, 1994).

Performance of sinusoidally deformed hydrophone line arrays

Deanna M. Caveny,[a] Donald R. Del Balzo, and James H. Leclere
Naval Research Laboratory, Stennis Space Center, Mississippi 39529

George E. Ioup
Department of Physics, University of New Orleans, New Orleans, Louisiana 70148

(Received 31 January 1997; revised 2 April 1998; accepted 14 December 1998)

It is well known that array deformations can distort beam patterns and introduce bearing errors if the beamformer assumes linearity. It is also known that deformed arrays can resolve left–right ambiguities, provided the shape is known. In this work, these two effects are studied for undamped and damped sinusoidally deformed arrays with small deformation amplitudes in the horizontal (x,y) plane only. By use of fixed arc-length separations along the array, the hydrophone (x,y) coordinates are determined numerically and the error in assuming equal x spacing is summarized for a sample array. Array-response patterns are analyzed for two conditions: (1) when the deformed array shape is assumed linear and (2) when the deformed array shape is known exactly. Degradations resulting from assuming linearity and the ability to resolve left–right ambiguities are discussed in terms of reduced gain, degraded angular resolution, and bearing errors. Shape-unknown signal-gain degradation ranges to 7 dB at broadside, but is less than 1 dB near endfire. For the shape-known case, signal gain for the true peak is greater than signal gain for the ambiguous peak by up to 9 dB for sources at broadside and to just over 2.5 dB for arrivals near endfire. © *1999 Acoustical Society of America.* [S0001-4966(99)06103-2]

PACS numbers: 43.30.Wi, 43.30.Bp [SAC-B]

INTRODUCTION

Hinich and Rule,[1] Hodgkiss,[2] Bouvet,[3] Ginzkey,[4] and Butler[5] have shown that deformations from a straight-line shape in the horizontal plane of towed arrays can produce significant distortions in array-response patterns and errors in bearing estimation if the beamformer assumes linearity. Hinich and Rule[1] use approximate undamped and damped sinusoidal shapes and report the case of $3\frac{1}{2}$ half-cycles of the sinusoid. For the damped case, deformation increases with distance from the towing platform. Hodgkiss[2] employs a single circular arc shape and discusses errors in passive ranging and bearing estimation. Bouvet[3] develops a model for large random array variations using fixed sensor separations (nonelastic array) with application to a circular arc. Bouvet[3] also gives a helpful brief review of related literature. Ginzkey[4] studies the effects of small two-dimensional random position errors. Butler[5] uses a sinusoidal deformation model which assumes equal x spacing of the hydrophones.

More recent work has discussed nonacoustic and acoustic methods to estimate array shapes, without emphasis on performance implications. One nonacoustic method involves direct hydrodynamic modeling based on single-point measurements either on the tow ship or on the cable itself with motion propagated along the array.[6–9] Another is based on distributed measurements from nonacoustic sensors along the array (e.g., depth gauges and compasses).[10,11] Generally, these techniques rely on solving the Paidoussis equation and/or interpolating between known points with polynomials or splines. The acoustic approaches involve a variety of signal-processing techniques using acoustic signals received at the hydrophones in two general categories—(a) from near-field controlled sources, and (b) from far-field noncontrolled sources of opportunity. The first approach usually involves arrival-time measurements from explosive sources[12,13] and the second often exploits relative phase information by working in the frequency domain.[14–16]

The issue of practical determination of array shapes is addressed well by the references above and others, and is not discussed further. The work reported here examines the impact (either good or bad) of array deformations (both known and unknown) in terms of beamformer performance and left–right ambiguity resolution.

This work examines the performance of towed arrays with small, horizontal deformations, primarily caused by unplanned variations in the tow-ship trajectory. The array-shape model is also capable of treating larger deformations, which could result from planned tow-ship maneuvers. The physical basis for the shape model derives from a harmonically driven damped oscillator, with small steering corrections of the towing platform providing the driving force. The attachment (or tow) point between the steel tow cable and the neutrally buoyant horizontal array is the origin for this model, and it is approximated to be a fixed node. A short vibration isolation module (VIM) is inserted between the tow point and the hydrophone array. The model produces an array shape based on the number of cycles (whole or fractional), the amplitude, and a damping factor. A drogue is assumed to be attached to the aft end of the array; thus, the damping in this model decreases the deformation as one moves away from the tow point, in contrast to the model of Hinich and Rule.[1]

This study is based on acoustic field modeling and

[a]Present address: Department of Mathematics, College of Charleston, Charleston, SC 29424.

beamforming using computer software[17] that generates cross-spectral matrices for arbitrary hydrophone locations in specified noise fields. The results presented here use conventional beamforming with infinitely high signal-to-noise ratio on horizontal arrays with sinusoidal deformations. The fixed arc-length method of defining array shapes is described, along with a comparison of beamformer performance between known and unknown shapes. Finally, a discussion of the impact of array deformation on gain and bearing ambiguity resolution is given.

I. DETERMINATION OF HYDROPHONE *X*– *Y* LOCATIONS

To approximate a sinusoidal shape, Hinich and Rule[1] use straight-line segments between hydrophones. To calculate the locations of hydrophones for the sinusoidal models without approximation, however, it is necessary to fix the hydrophone spacing along the array curve and determine the x and y (horizontal plane) coordinates. This models an elastic array with varying sensor separations, overcoming the limitation discussed by Bouvet.[3] The vertical variable z is assumed constant for this study. The method for determining the coordinates involves the numerical evaluation of the arc-length integral. In the limit of small sinusoidal amplitudes, the hydrophones can be assumed to have equally spaced x locations, greatly simplifying the calculation. For the present work, only the arc-length integral method is used, and neither the assumption of equally spaced x locations[5] nor the assumption of straight-line segments between hydrophones[1] is employed.

The natural dimensions for scaling position variables and other length measures for deformed, equally spaced hydrophone arrays are the array element spacing, d, and the design wavelength, λ, which is assumed to be $2d$. The array is simulated to contain a forward VIM with arc-length of $6d$, followed by 128 hydrophones, and terminated by a drogue for stability. The first hydrophone is at an arc-length of $d/2$ from the point where the VIM connects to the hydrophone array. Each succeeding sensor is separated by an arc-length d along the curve from the previous one.

The problem is stated as follows. Assume that the towed array takes the shape of an undamped or damped sinusoid. Given a specific number of cycles, the undamped amplitude, and the amount of damping, determine the (x,y) coordinate location of each hydrophone. An equation for the array shape can be written as

$$y(x) = Ae^{-ax}\sin(\pi x/w), \tag{1}$$

where the undamped amplitude, A, and the amount of damping, a, are specified. The third parameter, w, although fixed by the number of cycles, is not known initially. It is to be determined before the coordinates are calculated.

Consider an undamped sine curve of p cycles. Let L denote the total array length, which is $(N+5.5)\,d$ if there are N hydrophones and the VIM is $6d$ in length. Then, the arc-length between two adjacent nodes for an undamped array is $L/2p$. The arc-length integral is given by

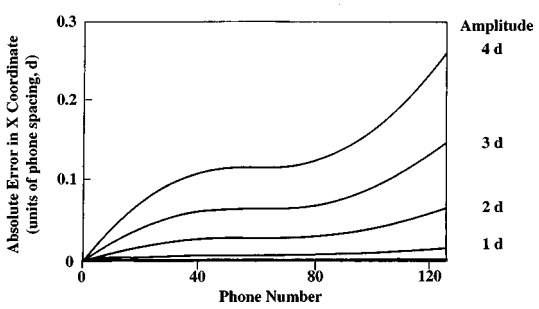


FIG. 1. The absolute error in the x-coordinate vs hydrophone number of a 128-element array resulting from the assumption of equal spacing d along the x axis. The array shape is an undamped 1/2-cycle sine, with deformation amplitudes from $0.5d$ to $4.0d$.

$$I_U(A,L) = \int_0^u [1 + A^2q^2\cos^2 qx]^{1/2}\,dx, \tag{2}$$

where $q = p/w$ and u and w are to be determined. Various approaches may be used; for example, one could choose $u = L/2p$ and $w = u$ and use numerical integration to refine these values until a specified tolerance between the calculated and known arc-lengths is reached.

In the damped case, the equal spacing of the zero-crossings (or nodes) is preserved, but the array length between any two adjacent nodes is no longer a constant. The arc-length integral is given by

$$I_D(A,L,a) = \int_0^u [1 + A^2e^{-2ax}(a\sin qx - q\cos qx)^2]^{1/2}\,dx. \tag{3}$$

In this case, the upper limit of the arc-length integral (i.e., the unknown value u) is chosen to be the x-coordinate of the last hydrophone. Then, the known arc-length is the total array length. Initially, u is taken to be L, and $w = u/2p$. The arc-length expression (3) can be evaluated and u adjusted, with $w = u/2p$, until the integral is close enough to L.

The x-coordinate of each hydrophone is found in a similar fashion, except that w is now determined and u gives the hydrophone x-coordinate. The integration arc-length is initially from the tow point to the first hydrophone, or generally from the last known hydrophone location to the adjacent unknown location. The corresponding y-coordinates are easily calculated from Eq. (1).

If instead one assumes that the x-coordinates are equally spaced with spacing d, the numerical integration could be avoided. For sine curves with small amplitudes, this assumption introduces only small errors. But the magnitude of the error grows with increasing hydrophone number and increasing array-deformation amplitudes. The assumption of this equal spacing always shifts the x-coordinates in a positive direction, making the array appear longer than it actually is, and the accumulated error increases more rapidly when the tangent line to the sine curve is steeper. Figure 1 illustrates the absolute value of the error in the x-coordinate of each hydrophone as a function of hydrophone number for arrays with 1/2-cycle distortion of various deformation amplitudes. The cumulative effect of the equal spacing assumption is evident, especially for the larger array amplitudes. The de-

TABLE I. Array geometries.

Cases	σ/λ	A	a	Shape
a	0.0	0.0	0	Linear
b	0.3	2.13	0	Half cycle
c	0.3	1.47	0	Full cycle
d	0.3	0.87	0	1 1/2 cycle
e	0.2	2.13	0.0069	Half cycle
f	0.1	2.13	0.0200	Half cycle

viations of the true x positions from equal x spacing do not become larger than $0.1d$ ($\lambda/20$) until the deformation of the array is greater than $2d$ for a 1/2-cycle sine array of 128 hydrophones.

II. EXAMPLES OF DEFORMED ARRAY BEAMFORMING

Hodgkiss[2] investigates plane-wave beamforming for various source locations and circular arc array shapes. His results are given as array-response plots when beamforming with both the actual circular arc hydrophone locations and assumed linear locations. He does not consider left–right ambiguity resolution and his array-response patterns go over only 180 deg. Similar studies are conducted here for arrays having undamped and damped sinusoidal geometries, with the addition of an examination of left–right ambiguity resolution and the calculation of performance curves.

Six array geometries are considered in this study: (a) a linear array for reference; (b) an undamped 1/2-cycle deformation with amplitude of 2.13 hydrophone spacings; (c) an undamped full-cycle deformation with amplitude of 1.47 hydrophone spacings; (d) an undamped 1 1/2-cycle deformation with amplitude of 0.87 hydrophone spacings; (e) a damped 1/2-cycle deformation with maximum amplitude of 1.55 hydrophone spacings ($A=2.13d$ and $a=0.0069$); and (f) a more highly damped 1/2-cycle deformation with maximum amplitude of 0.95 hydrophone spacings ($A=2.13d$ and $a=0.020$). These amplitude and damping factor values were chosen to produce a value for the undamped cases of 0.3 in the array shape statistic, σ/λ, with σ the rms shape distortion as measured from a best-fitting straight line, and values of 0.2 and 0.1, respectively, for the damped cases. The cases are summarized in Table I.

The source azimuths considered in this section are 90 (broadside), 45, and 10 deg from endfire, all at the design frequency and all in the horizontal plane. Calculations for out-of-plane arrivals (10 deg from the horizontal) were made and shown to be consistent with the in-plane results (to within 0.003 dB) and therefore are excluded from the study. Figure 2 illustrates the beamformed array-response patterns (with equal weighting on each hydrophone and no background noise) for a linear array over the full 360 deg azimuthal sector. The upper plot shows the 90 deg (broadside) source azimuth result. The middle and lower plots show the 45 and 10 deg source results, respectively. Note the standard results of beam broadening away from broadside and the occurrence of grating lobes as the signal approaches endfire. Figures 3–5 contain array-response patterns for sinusoidally deformed arrays assuming that beamforming is implemented

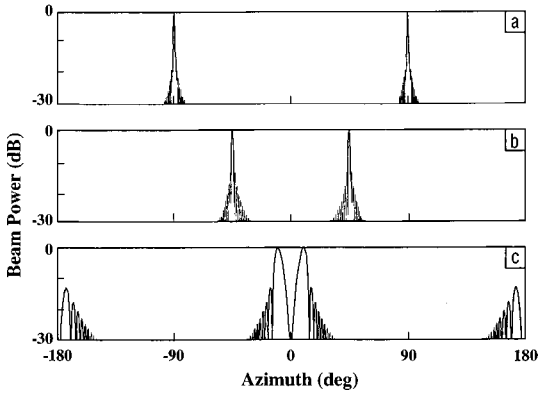


FIG. 2. Linear array responses for sources at (a) 90 deg, (b) 45 deg, and (c) 10 deg from endfire.

with both (1) the incorrect assumption that the array geometry is linear, and (2) the actual hydrophone locations known. The responses shown in Fig. 2 are included so that the deformed array responses can be compared. Beam powers for all figures are referenced to 0 dB for the linear array-response maximum at a given source direction. None of the responses below -30 dB is plotted.

Figure 3 shows the response of the undamped 1/2-cycle sine array with $\sigma/\lambda=0.3$ to sources at 90 deg in (a) and (b), 45 deg in (c) and (d), and 10 deg in (e) and (f). In Fig. 3(a), (c), and (e), the array shape is assumed known and the actual element locations are used in the beamforming. Since the distorted array has almost the same total aperture as the linear array, the forward (true) peak is almost identical to that for the linear response. The ambiguous (false) peak, however, does not have the same phase delays for the deformed array as the forward peak does, so it is significantly changed. It has less signal gain, is broader, and is broken up into several local maxima for the sources at 90 and 45 deg. While the ambiguous peak at -10 deg (corresponding to a source at 10 deg) is somewhat reduced and broadened, it is not broken up in the same way as the others. This is due to two factors: (a) the array has less resolution (wider beams) near endfire than at broadside, and (b) a plane wave arriving in a direction close to endfire sees a smaller array deformation than one arriving at broadside. If, as is generally the case, the

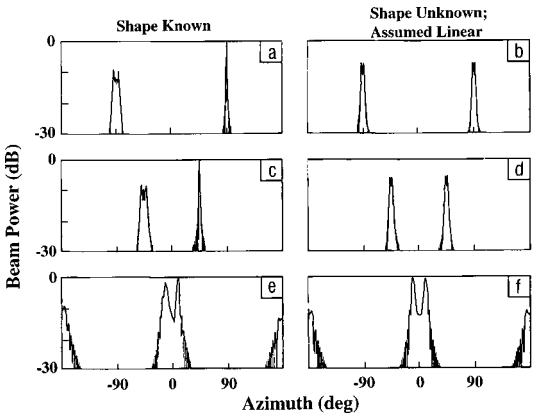


FIG. 3. Undamped 1/2-cycle deformed-array response. Amplitude of deformation is $2.13d$ and σ/λ is 0.3. Source is at 90 deg for (a) and (b), at 45 deg for (c) and (d), and at 10 deg for (e) and (f).

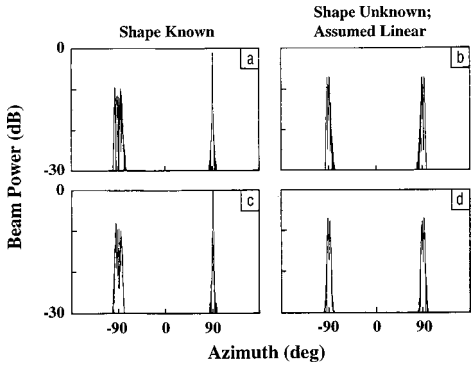


FIG. 4. Undamped full-cycle, (a) and (b), and 1 1/2-cycle, (c) and (d), deformed-array responses for a source at 90 deg. Deformation amplitude for full cycle is $1.47d$, and for 1 1/2 cycles is $0.87d$. σ/λ for both cases is 0.3.

array shape is unknown and beamforming is done assuming the shape to be linear, the responses of Fig. 3(b), (d), and (f) result. The signal gain is reduced, especially at 90 and 45 deg, where the response peaks are also split. At 10 deg, the reduction in gain is small and the main peak shape is close to that of the linear response, again because the deformation looks smaller and the beams are wider near endfire.

For the remaining array shapes, only the array response to a broadside arrival is shown. The second and third undamped examples are in Fig. 4, while the damped cases are illustrated in Fig. 5. For the undamped arrays, $\sigma/\lambda=0.3$, the same value as the 1/2-cycle undamped array of Fig. 3. The general behavior of the responses of the full-cycle array, Fig. 4(a) and (b), and the 1 1/2-cycle array, Fig. 4(c) and (d), is similar to that of the broadside responses of the 1/2-cycle array. The ambiguous peak in the shape-known responses and both the true and ambiguous peaks in the shape-unknown responses exhibit fine structure. This is because the deformed-array shapes themselves have structure. In effect, the deformed array is composed of several nearly straight subsections, each of which has its own natural direction. Thus, the incident plane wave is resolved into multiple directions.

Since the damped arrays of Fig. 5 (both 1/2 cycle) have smaller values of σ/λ (0.2 and 0.1) than the undamped cases, the shape-unknown responses are closer to the linear array

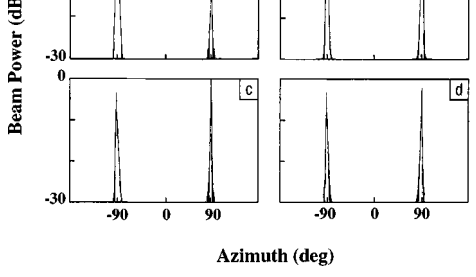


FIG. 5. Damped 1/2-cycle deformed-array response for a source at 90 deg. (a) and (b) are for a shape determined by $A=2.13d$, $a=0.0069$, and $\sigma/\lambda=0.2$, while (c) and (d) are for an array with $A=2.13d$, $a=0.020$, and $\sigma/\lambda=0.1$.

response at broadside than the undamped responses of Figs. 3 and 4. Thus, for the unknown-shape case, the drogue seems beneficial to performance because it increases damping, which in turn decreases physical deformations, leading to reduced beamformer phase-delay errors when linearity is assumed. This produces greater signal power through the beamformer.

The irregular nature of the broken peaks in Figs. 3 through 5 leads to instabilities in such performance measures as peak height, bearing, and beamwidth because of the difficulty in defining these quantities. The splitting of the true peak when the beamforming is done assuming a linear array leads to bearing errors resulting from choosing the largest subpeak. This suggests that for arrays which have a large enough aperture and enough deformation to produce this splitting, it may be better to fit a smooth analytic shape in order to estimate signal gain, source direction, and beamwidth.

Note that the shape-unknown responses are all symmetric about 0 deg in Figs. 3–5. This is because differences in the field as sensed by a distorted array correspond to phase shifts (from the phases of a linear array) that are equal and opposite to the phase errors in the steering vectors that result from assuming that the distorted array is straight. To understand this result, consider the phases at the hydrophones for arrival directions of plus and minus θ. For the deformed arrays, arrivals from $+\theta$ will have, at each hydrophone, a shifted phase Δ_+ from the phase value at a straight line array, and arrivals from $-\theta$ will have a different shifted phase Δ_-. These phase shifts will be incorporated into the cross-spectral matrix for both shape-known and shape-unknown beamforming. For shape-unknown beamforming, the steering vectors correspond to a linear array. Thus, the phase errors in these steering vectors are opposite to the phase differences in the cross-spectral matrix mentioned above, and therefore the plus and minus arrival directions have the same (incorrect) array response.

III. PERFORMANCE DEGRADATION FOR DEFORMED ARRAYS

Hodgkiss[2] quantifies degradations in the beamforming process, with the incorrect assumption of linearity, for known circular arc shapes as a function of the amount of bow. This section contains a systematic study of performance degradation for sinusoidally deformed arrays when the shape is unknown, in terms of three measures: (1) signal gain, (2) beamwidth broadening, and (3) bearing shifts. The signal-to-noise ratio is infinite and the signal degradation is considered for σ/λ in the range of 0.0 to 0.3.

Figure 6 addresses the first issue, signal gain, by showing the power loss in the true peaks in the shape-unknown case, relative to the linear-array peak power, plotted versus σ/λ, for various array damped and undamped shapes (half cycles, full cycles, 1 1/2 cycles) and for various source azimuths. The azimuths selected are 10, 30, 45, 60, and 90 deg for the 1/2-cycle cases, and 10 and 90 deg for the others. For the damped array, σ/λ may be varied by changing either the amplitude (controlled primarily by tow-ship trajectory varia-

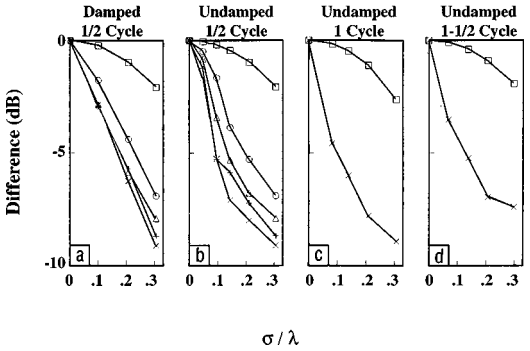


FIG. 6. Performance curves of peak-height degradation for deformed arrays assumed to be linear. Loss in array-signal gain of deformed arrays is given as the power loss in the source peaks, in dB, relative to the corresponding peaks for a linear array, versus σ/λ for sources at 10 (□), 30 (○), 45 (△), 60 (+), and 90 deg (×).

tions) or the damping factor (controlled primarily by the drogue characteristics). This performance evaluation is conducted by varying the damping factor.

For all cases, the degradation is greatest for broadside arrivals (6–7 dB when $\sigma/\lambda=0.3$), decreasing, in general, as the source arrival angle approaches endfire (about 0.5 dB when $\sigma/\lambda=0.3$ for the 10 deg source). The small exceptions, as before, are due to the irregular qualities of the degraded peaks. As expected, the degradation becomes worse with increasing array deformations (i.e., greater σ/λ) for all shapes. For a given source direction, the degradations are similar for all combinations considered, except for the results corresponding to the 90 and 60 deg source directions. For these arrival angles, among the cases examined, only the 1/2-cycle undamped and the full-cycle performance for sources at broadside track fairly closely. The performance at $\sigma/\lambda=0.3$, however, is identical for all arrival angles for the damped and undamped 1/2-cycle cases because the two array shapes are identical since the damped array has $\sigma/\lambda=0.3$ when the damping is exactly zero.

One practical application of performance summaries, such as those shown in Fig. 6, is to determine, as a function of σ/λ, if the array-element locations need be known or if the beamforming process can assume a linear array. As an example, given a full-cycle, damped or undamped, deformed array and broadside arrivals, if no more than a 5-dB loss in signal gain is acceptable, then array-element locations are needed when $\sigma/\lambda>0.2$. If no more than a 3-dB loss is tolerable, then the approximate upper limit for assuming linearity is $\sigma/\lambda=0.13$. These findings are consistent with the general loss in signal gain for Gaussian errors in element locations given by Steinberg[18] in his Fig. 6 and the accompanying discussion. Note that these σ/λ limits are a function of array shape, and that for broadside arrivals they are higher for 1 1/2-cycle arrays and lower for damped and undamped 1/2-cycle arrays.

Second, distorted arrays can produce beam broadening. One can consider the true-peak beamwidth for shape-unknown beamforming as a measure of performance degradation by comparing it to the beamwidth for the true peak in the corresponding linear-array response. Although not quantified here, significant true-peak broadening can be observed in the shape-unknown response patterns of Figs. 3 through 5. Performance curves for true-peak broadening as a result of assuming linearity serve as a measure, which, along with the loss in array-signal gain, can be used to determine the largest acceptable value of σ/λ for shape-unknown beamforming.

Third, distorted arrays can produce bearing errors. For small values of σ/λ, incorrectly assuming a linear array may result in only small losses in signal gain and beam resolution. In these instances, one may choose to accept this degradation. As Hinich and Rule[1] and Hodgkiss[2] point out, however, there can still be a bearing error of 1 to 2 deg. This bearing error arises from the splitting of the true peak into two or more subpeaks, the largest of which is not centered with respect to the peak spread. For the deformed-array responses shown in this paper, only damped 1/2-cycle responses are included for deformations with σ/λ less than 0.3. In Fig. 5(b), $\sigma/\lambda=0.1$ and the peak is already asymmetrical, although not highly broken. For $\sigma/\lambda=0.2$, the response shown in Fig. 5(d) is split into two parts with a minimum between them at the correct source bearing. The broadside $\sigma/\lambda=0.3$ peaks, shown for various array shapes in Figs. 3(b) and 4(b) and (d), exhibit behavior ranging from a simple splitting into two parts to a highly broken and irregular shape. Thus, it is understandable that even relatively small array deformations lead to bearing errors as large as approximately half the true-peak beamwidth in shape-unknown beamforming. Hinich[19] and Bouvet[3] (and references cited therein) discuss techniques for estimating the correct bearing.

IV. LEFT–RIGHT AMBIGUITY RESOLUTION FOR DEFORMED ARRAYS WITH KNOWN SHAPE

The standard technique to resolve left–right ambiguities on nominally straight towed arrays is first to record the two possible true bearings toward a source, second to make a course change, and third to note the new possible true bearings. A consistency check will give the correct bearing. Unfortunately, during a course change, uncorrected array deformations can be so severe that loss in beamformer signal gain can cause a loss in source detection (against noise). After the turn is completed and the tow-ship trajectory has stabilized, there is still a residual time required for the array to straighten and stabilize. For some applications, these time delays are unsatisfactory. A process which could allow continuous monitoring of the true source bearing without loss in detection time is desirable.

Both Hinich and Rule,[1] and Hodgkiss[2] discuss advantages of a deformed array over a linear array to discriminate true from ambiguous peaks. This section examines two approaches for continuous left–right ambiguity resolution for sinusoidally deformed arrays when the shape is known. The first involves the power difference, and the second involves the beam width ratio between the true and ambiguous peaks. All of the results are discussed in terms of the amount of array distortion, as defined by the σ/λ measure, with infinite signal-to-noise ratio. Figure 7 illustrates the power loss in the ''false,'' or ambiguous, peak. This loss is plotted vs σ/λ for various array damped and undamped shapes (half cycles, full cycles, 1 1/2 cycles) and for various source azimuths. The

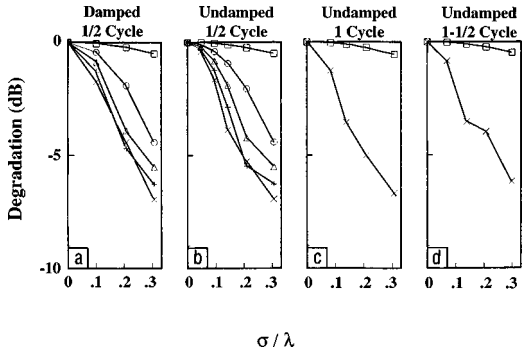


FIG. 7. Performance curves for peak-height difference in left–right ambiguity resolution of deformed arrays with known element locations. Amount by which left (ambiguous) peak is down from right (true) peak is given in dB versus σ/λ for sources at 10 (□), 30 (○), 45 (△), 60 (+), and 90 deg (×).

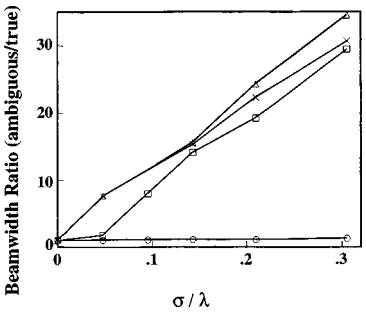


FIG. 8. Left–right ambiguity beamwidth resolution. The beamwidth ratio, defined as the false-peak 3-dB width over the true-peak 3-dB width, for deformed arrays beamformed with array-element locations known, versus σ/λ. Results are given for a 1/2-cycle array (□), a full-cycle array (△), and a 1 1/2-cycle array (×) for a broadside source and also for a 1/2-cycle array (○) with a source at 10 deg.

azimuths selected are 10, 30, 45, 60, and 90 deg for the 1/2-cycle cases and 10 and 90 deg for the others.

In Fig. 7(a), the 1/2-cycle damped array performance is given for left–right ambiguity resolution in terms of true peak minus ambiguous peak power difference in dB versus σ/λ. In general, for this and all cases in Fig. 7, the ability to discriminate an ambiguous peak from a true peak by power difference is greatest for sources at broadside, and decreases to be least for sources close to endfire. This trend is expected because the left–right phase difference is smaller at endfire. Deviations from this observation are slight in Fig. 7, and occur because of the instabilities in the broken ambiguous peak maxima discussed earlier.

The first observation from Fig. 7 is based on a comparison of 7(a) and (b) where the array is distorted into the same general shape (i.e., 1/2 cycle) but with and without damping. For a given σ/λ, the undamped array almost always has greater power differences, and is therefore a better left–right source discriminator, for source azimuths away from endfire. Thus, damping is generally deleterious to performance when attempting to resolve left–right ambiguities by true peak–ambiguous peak power differences. This is in contrast to the previous conclusion that damping is beneficial when considering beam power (signal gain) on a distorted array assumed linear. There is a tradeoff between the two countering effects which can be evaluated for a given scenario.

In Fig. 7(b), (c), and (d), the left–right ambiguity resolution performance for the undamped 1/2-cycle, full cycle, and 1 1/2-cycle arrays can be compared. At broadside, for a given σ/λ, the undamped half and full cycle arrays are better (i.e., have greater power difference) at resolving left–right ambiguity than a 1 1/2-cycle array. An examination of Figs. 3(a) and 4(a) and (c), however, shows that for $\sigma/\lambda=0.3$, this advantage in ambiguity resolution is due mainly to two thin spikes in a highly broken 1 1/2-cycle ambiguous peak. If an average or curve-fit peak is used instead of the tallest subpeak to measure ambiguity resolution, this distinction in the difference performance measure is not expected to be as large.

The other approach for left–right ambiguity resolution concerns beam broadening. Beamforming with the known hydrophone locations gives true peaks which correspond closely, in terms of 3-dB beamwidth, to the true peaks which a linear array would produce for the small deformations considered here. For the ambiguous peaks, on the other hand, the broadening is large and the ratio of the 3-dB beamwidth of the ambiguous peak to that of the true peak may be taken as another measure of left–right ambiguity resolution. In Fig. 8, this ratio is given, as a function of σ/λ, for all three undamped cases with broadside arrivals and for the undamped 1/2-cycle case with a 10 deg arrival. For broadside incidence, the ratio of the beamwidths for the ambiguous peak to the true peak increases rapidly with increasing σ/λ to a value of 30 to 35 at $\sigma/\lambda=0.3$ for all three shapes. It is possible that, at small deformations, this ratio may be a better discriminator for left–right ambiguity resolution than the difference in signal gain for sources at broadside in some applications. For the 1/2-cycle deformation and a source at 10 deg, however, the beamwidth ratio is almost constant at 1, versus σ/λ, and so would not serve as a useful discriminant. The 3-dB beamwidths of the broken ambiguous peaks have been determined as accurately as possible without recourse to curve fitting and may be subject to small errors.

It should be noted that the shape-unknown beamwidths in the previous section are smaller for all these examples than the beamwidths of the ambiguous peaks in the associated shape-known responses. Therefore, in this limited range of calculations for shape-known beamforming, the ratios shown in Fig. 8 are larger than would be found for shape-unknown beamforming. This result is not surprising, since the phase errors for the ambiguous peaks for shape-known beamforming are, in a sense, twice those of shape-unknown beamforming.

V. CONCLUSIONS

This paper reports results of array performance as affected by known and unknown distortions in array shape. Using a simple but accurate model of hydrophone positions, which produces an array with equal arc-lengths between elements, various array configurations were constructed. These included undamped and damped 1/2-cycle sinusoidal configurations and also undamped full-cycle and 1 1/2-cycle configurations.

In shape-known beamforming, the ability to discriminate true peaks from ambiguous peaks increases as array deformation increases. Differences in array-signal gain for these two peaks range up to 9 dB for $\sigma/\lambda = 0.3$ when the source is at broadside. For arrivals near endfire, however, the largest difference is only about 2.5 dB. In shape-unknown beamforming, the degradation in array-signal gain ranges up to 7 dB at broadside, but remains less than 1 dB near endfire. The results can be used to determine if shape estimation is required.

Beamwidths of the ambiguous peaks were compared to beamwidths of the true peaks in shape-known beamforming. The ratio of these beamwidths increases rapidly with σ/λ, reaching a value of 30–35 at $\sigma/\lambda = 0.3$ for the full-cycle array with the source at broadside. The true-peak broadening for shape-unknown beamforming is also significant, but less than that of the ambiguous peak. Near endfire, the broadening is negligible for both types of beamforming for the domain of σ/λ considered.

Straightforward measures of array signal-gain degradation and beamwidths are difficult to apply due to the broken nature of the peaks with the resolution capability of 128 hydrophones (64-λ array). This problem also leads to errors in bearing estimation.

Regarding the question of the utility of drogues to stabilize towed arrays, there is an apparent dichotomy. Drogues reduce array horizontal deformations, and this improves signal gain for the shape-unknown case with linearity assumed. However, array straightening hinders left–right signal discrimination. Thus, use of drogues may depend on the specific objectives and scenarios.

ACKNOWLEDGMENTS

The authors wish to acknowledge the financial support of the Naval Undersea Warfare Center and of the Office of Naval Research. The authors are extremely grateful to Art Collier of the Canadian Defense Research Establishment Atlantic for supplying BEAMSTATPAK, a powerful and general computer-software package for modeling acoustic signal and noise fields and for beamforming. Helpful discussions with Don Murphy and Jeff Beckleheimer of NRL are also acknowledged.

[1] M. J. Hinich and W. Rule, ''Bearing estimation using a towed array,'' J. Acoust. Soc. Am. **58**, 1023–1029 (1975).
[2] W. S. Hodgkiss, ''The effects of array shape perturbation on beamforming and passive ranging,'' IEEE J. Ocean Eng. **8**, 120–130 (1983).
[3] M. Bouvet, ''Beamforming of a distorted line array in the presence of uncertainties on the sensor positions,'' J. Acoust. Soc. Am. **81**, 1833–1840 (1987).
[4] L. Ginzkey, ''Influence of sensor position errors on spatial signal processing algorithms,'' in *Proceedings of NATO ASI on Adaptive Methods in Underwater Acoustics*, Luneburg, Germany, 30 July–10 August 1985, edited by H. Urban (Reidel, Dordrecht, 1985), pp. 477–482.
[5] D. Butler, ''Beamforming with a distorted towed array,'' in *Proceedings of* NATO ASI *on Adaptive Methods in Underwater Acoustics*, Luneburg, Germany, 30 July–10 August 1985, edited by H. Urban (Reidel, Dordrecht, 1985), pp. 469–475.
[6] C. Lee, ''A modeling study on steady-state and transverse dynamic motion of a towed array system,'' IEEE J. Ocean Eng. **3**, 14–21 (1978).
[7] J. Ketchman, ''Vibration induced in towed linear underwater array cables,'' IEEE J. Ocean Eng. **6**, 77–86 (1981).
[8] D. R. Del Balzo, ''Observations of towed array vertical stability and horizontal shape during SACLANTCEN Cruise LRG 1/89,'' SACLANTCEN SR-220, La Spezia, Italy (1993).
[9] C. S. van Aartsen, ''The Response Model,'' in *Proceedings of Low Frequency Active Acoustics Conference*, La Spezia, Italy, SACLANT Undersea Research Center (1993).
[10] D. A. Gray, B. D. O. Anderson, and R. R. Bitmead, ''Towed array shape estimation using Kalman filters—theoretical models,'' IEEE J. Ocean Eng. **18**, 543–556 (1993).
[11] B. E. Howard and J. M. Syck, ''Calculation of the shape of a towed underwater acoustic array,'' IEEE J. Ocean Eng. **17**, 193–203 (1993).
[12] L. J. Rosenblum and D. R. Del Balzo, ''A constructive algorithm for determining array shape with application to Freddex,'' Naval Research Lab, Washington, DC, NRL Repert 8531 (1979).
[13] E. C. Ballegooijen, G. W. M. Mierlo, C. van Schooneveld, P. P. M. van der Zalm, A. T. Parsons, and N. H. Field, ''Measurement of towed array position, shape, and attitude,'' IEEE J. Ocean Eng. **14**, 375–383 (1989).
[14] B. G. Quinn, R. F. Barrett, P. J. Kootsookos, and S. J. Searle, ''The estimation of the shape of an array using a hidden Markov model,'' IEEE J. Ocean Eng. **18**, 557–564 (1993).
[15] B. G. Ferguson, ''Remedying the effects of array shape distortion on the spatial filtering of acoustic data from a line array of hydrophones,'' IEEE J. Ocean Eng. **18**, 565–571 (1993).
[16] D. E. Wahl, ''Towed array shape estimation using frequency-wavenumber data,'' IEEE J. Ocean Eng. **18**, 582–590 (1993).
[17] J. H. Leclere, D. R. Del Balzo, H. A. Chandler, R. R Slater, and G. E. Ioup, ''A brief introduction to BEAMSTATPAK with sample calculations of array performance on multiple line systems,'' NORDA Tech Note 49, Naval Res. Lab., Stennis Space Center, MS (1991).
[18] B. D. Steinberg, *Principles of Aperture and Array System Design* (Wiley, New York, 1976).
[19] M. J. Hinich, ''Bearing estimation using a perturbed linear array,'' J. Acoust. Soc. Am. **61**, 1540–1544 (1977).

Phenomenological theory of the translational relaxation times in gases

Allan J. Zuckerwar
Mail Stop 236, NASA Langley Research Center, Hampton, Virginia 23681

(Received 19 May 1998; revised 7 January 1999; accepted 9 January 1999)

The exact solution to the classical equations governing the translational dispersion and absorption of sound in a gas obscures its relaxational character because of its mathematical complexity. The approach taken here is to solve the secular equation by the method of Padé approximants, which even to the relatively low order R_{11} yields a remarkably close approximation to the exact solution over a wide range of frequency/pressure (f/P) ratios. As a result, translational relaxation can be formulated in terms of a conventional relaxation process with well-defined relaxation times, relaxation strength, collision numbers, additivity relations, etc. To extend the theory to high values of f/P ratio, a model is proposed to account for the noncontinuum behavior of the transport coefficients (viscosity and thermal conductivity) as the molecular mean free path approaches the acoustical enclosure dimensions. The theoretical dispersion and absorption show good agreement with measurements in argon over the classical and transition regions of f/P, but a discrepancy appears at higher values of f/P, where collective propagating modes, assumed in the theory, give way to single-particle modes, prevailing in the experiments. [S0001-4966(99)01704-X]

PACS numbers: 43.35.Ae, 43.35.Fj, 43.20.Hq [HEB]

INTRODUCTION

The classical equations governing the translational dispersion and absorption of sound in a gas yield a solution of such complexity as to obscure the relaxational character of the associated viscous and thermal transport processes. In the past, only in the case of decoupled viscous and thermal transport have the solutions been shown to take the form of a conventional relaxation process,[1] but these solutions are physically unrealistic because the viscous and thermal motions in acoustical propagation inherently are strongly coupled. In this paper a solution is presented which provides a well-defined relaxation strength and relaxation times for the strongly coupled viscous and thermal transport processes responsible for translational relaxation. The approach taken here is to solve the secular equation by the method of Padé approximants, which even to the relatively low order R_{11} yields a remarkably close approximation to the exact solution over a wide range of frequency/pressure ratios (or, equivalently, viscous Reynolds numbers), and is not encumbered by the convergence difficulties encountered in the solutions based on power series expansions.

As the frequency/pressure ratio increases to the point where the molecular mean free path becomes comparable to the acoustical enclosure dimensions, then the transport coefficients are no longer independent of pressure and the values used in continuum theory are no longer valid. As an alternative to analytical treatments based on the Boltzmann transport equation, the present theory retains the classical continuum equations but utilizes a proposed model to account for the pressure dependence of the transport coefficients. The result shows good agreement with measurement not only over the continuum region but also up to the transition to the free (collisionless) propagation region.

The following sections of this paper will (I) review the exact solution to the secular equation (based on Navier–Stokes), (II) present a new approximate solution based on the Padé R_{11} approximant, revealing the true relaxational character of the translational processes, and (III) introduce a new method of analyzing the transition to high-frequency/pressure ratios. It will be shown that translational relaxation can be ascribed the attributes of a conventional relaxation process like relaxation strength, relaxation times, collision numbers, additivity relations in the presence of other relaxation processes, etc. Application to monatomic gases, polyatomic gases, and liquids will be discussed.

I. REVIEW OF THE CLASSICAL SOLUTION BASED ON NAVIER–STOKES

Consider a one-dimensional, small-amplitude, harmonic sound wave propagating in a gas. The equations of state (ideal), continuity, Navier–Stokes, and thermal transport are written accordingly

$$P=\frac{\rho RT}{M}, \tag{1}$$

$$\frac{\partial u}{\partial x}+\frac{1}{\rho_0}\frac{\partial \rho}{\partial t}=0, \tag{2}$$

$$\rho_0\frac{\partial u}{\partial t}=-\frac{\partial P}{\partial x}+\frac{4}{3}\mu\frac{\partial^2 u}{\partial x^2}, \tag{3}$$

$$C_V\frac{\partial T}{\partial t}=\frac{P_0}{\rho_0^2}\frac{\partial \rho}{\partial t}+\frac{\kappa}{\rho_0}\frac{\partial^2 T}{\partial x^2}, \tag{4}$$

in which P, ρ, T, M, u, μ, κ, and C_V are the pressure, density, temperature, molecular mass, particle velocity, absolute viscosity, thermal conductivity, and specific heat at constant volume of the gas. Further, t and x are time and propagation

TABLE I. Notation of various authors for the viscous and thermal relaxation times.

Author(s)	Viscous	Thermal	Reference	Author's Eq(s).
This work	τ_ν	τ_h		(6, 7)
Bhatia	ω_ν^{-1}	ω_χ^{-1}	1	(4.3.5, 5.4.11)
Greenspan	$(4/3)(\omega\gamma_0 r)^{-1}$	$(\omega\gamma_0^2 s)^{-1}$	2	(5, 6)
Hunt	τ_ν	τ_κ	3	(3c-85)
Truesdell	X/ω	XY/ω	4	(3.2)
Markham, Beyer, and Lindsay	ω_ν^{-1}	ω_c^{-1}	5	(3.8, 4.9)
Morse and Ingard	$(4/3)(\ell_v/c)$	ℓ_h/c	6	(6.4.20)
Pierce	$\epsilon_\mu/i\omega$	$\epsilon_\kappa/i\omega$	7	(10.34)
Trusler	τ_v	τ_h	8	(2.3.28, 2.3.29)

distance, R is the universal gas constant, and P_0 and ρ_0 are ambient values in the absence of sound. Complex quantities will be written in boldface. For a harmonic wave of the form $\exp[i(\omega t-\boldsymbol{k}x)]$, Eqs. (1)–(4) yield the following secular equation:

$$\boldsymbol{X}^4(-i\omega\tau_h+\omega^2\gamma\tau_v\tau_h)+\boldsymbol{X}^2[1+i\omega(\tau_v+\gamma\tau_h)]-1=0, \quad (5)$$

where $\boldsymbol{X}=c_0/\boldsymbol{c}$ is the ratio of the relaxed sound speed (in the limit of zero frequency) to the complex sound speed, ω and $\boldsymbol{k}=\omega/\boldsymbol{c}$ are the acoustic angular frequency and complex wave number, $\gamma=C_P/C_V$ is the ratio of the specific heats at constant pressure and constant volume, and τ_v and τ_h are the viscous and thermal relaxation times defined as

$$\tau_v=\frac{4}{3}\frac{\mu}{\rho_0 c_0^2} \quad (6)$$

and

$$\tau_h=\frac{\kappa}{\rho_0 c_0^2 C_P}. \quad (7)$$

Equation (5) agrees with Greenspan[2] if one replaces his k with our $-i\boldsymbol{k}/k_0$, where $k_0=\omega/c_0$, and agrees with Hunt[3] if one replaces his $\alpha+ik$ with our $i\boldsymbol{k}$. Comparable expressions for τ_v and τ_h in the notation of other authors are listed in Table I.[1–8]

The exact solutions to Eq. (5) are expressed in terms of the real and imaginary parts of $(c_0/\boldsymbol{c})^2$:

$$\mathrm{Re}\left[\left(\frac{c_0}{\boldsymbol{c}}\right)^2\right]=\frac{B^{1/2}}{2\omega\tau_h}\sin(\vartheta_2)[1\pm A^{1/4}\cos(\vartheta_1)] \pm\frac{B^{1/2}A^{1/4}}{2\omega\tau_h}\cos(\vartheta_2)\sin(\vartheta_1), \quad (8)$$

$$\mathrm{Im}\left[\left(\frac{c_0}{\boldsymbol{c}}\right)^2\right]=-\frac{B^{1/2}}{2\omega\tau_h}\cos(\vartheta_2)[1\pm A^{1/4}\cos(\vartheta_1)] \pm\frac{B^{1/2}A^{1/4}}{2\omega\tau_h}\sin(\vartheta_2)\sin(\vartheta_1), \quad (9)$$

where

$$A=\frac{[1-\omega^2(\tau_v-\gamma\tau_h)^2]^2+4\omega^2[\tau_v+(\gamma-2)\tau_h]^2}{[1-\omega^2(\tau_v+\gamma\tau_h)^2]^2+4\omega^2(\tau_v+\gamma\tau_h)^2}, \quad (10)$$

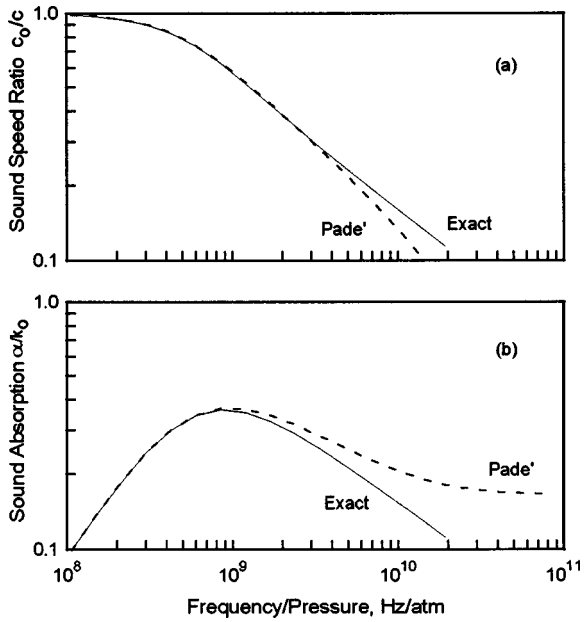


FIG. 1. Classical (a) dispersion and (b) absorption of sound in argon at 20 °C versus frequency/pressure ratio. Solid lines: exact theory, Eqs. (14) and (15). Dashed lines: approximate (Padé) theory, Eqs. (20) and (21).

$$B=\frac{1+\omega^2(\tau_v+\gamma\tau_h)^2}{1+\omega^2\gamma^2\tau_v^2}, \quad (11)$$

$$\vartheta_1=\frac{1}{2}\tan^{-1}\left\{\frac{2\omega[\tau_v+(\gamma-2)\tau_h]}{1-\omega^2(\tau_v-\gamma\tau_h)^2}\right\} -\frac{1}{2}\tan^{-1}\left\{\frac{2\omega(\tau_v+\gamma\tau_h)}{1-\omega^2(\tau_v+\gamma\tau_h)^2}\right\}, \quad (12)$$

$$\vartheta_2=\tan^{-1}[\omega(\tau_v+\gamma\tau_h)]-\tan^{-1}(\omega\gamma\tau_v). \quad (13)$$

Retaining the lower (negative) sign for acoustical wave propagation, one finds the sound speed ratio and absorption from Eqs. (8) and (9), the latter expressed in terms of the normalized absorption coefficient α/k_0:

$$\frac{c_0}{c}=\frac{1}{\sqrt{2}}\left\{\left[\mathrm{Re}\left(\frac{c_0^2}{\boldsymbol{c}^2}\right)\right]^{1/2}+\left[\mathrm{Re}\left(\frac{c_0^2}{\boldsymbol{c}^2}\right)+\mathrm{Im}\left(\frac{c_0^2}{\boldsymbol{c}^2}\right)\right]^{1/2}\right\}^{1/2}, \quad (14)$$

$$\frac{\alpha}{k_0}=\frac{1}{\sqrt{2}}\left\{-\left[\mathrm{Re}\left(\frac{c_0^2}{\boldsymbol{c}^2}\right)\right]^{1/2}+\left[\mathrm{Re}\left(\frac{c_0^2}{\boldsymbol{c}^2}\right)+\mathrm{Im}\left(\frac{c_0^2}{\boldsymbol{c}^2}\right)\right]^{1/2}\right\}^{1/2}, \quad (15)$$

where α is the absorption coefficient per unit distance, defined by $\boldsymbol{k}=k-i\alpha$. The sheer complexity of Eqs. (8)–(15) conceals their relaxational character; rather this is brought out by plots of these equations for argon at 20 °C, shown as solid lines in Figs. 1(a) and 1(b), which reveal the familiar dispersion step (more apparent on linear sound speed and frequency scales) and absorption peak. There are two characteristics of translational relaxation which are not shared by the more common relaxations of the internal degrees of freedom of a polyatomic molecule. First, an analytical form of the translational relaxation times, which is needed to predict the halfway point of the dispersion curve and the location of the absorption peak, would be difficult to derive from Eqs. (8) to (15), and to the author's knowledge, such a derivation has never been attempted. Second, the relaxation strength is

unity, as may be found from the zero-frequency and infinite-frequency limits of Eq. (8). The latter, being zero, confirms the conclusion that the sound speed ratio [reciprocal of Eq. (14)] increases without limit with increasing frequency/pressure, as was pointed out by Truesdell.[4]

II. APPROXIMATE SOLUTION BASED ON PADÉ APPROXIMANT R_{11}

The Padé method assumes a solution in the form of the quotient of two polynomials.[9] The R_{11} approximant, truncating the numerator and denominator after the first two terms, is written here as

$$X^2 \approx R_{11}(\omega) = \frac{1 + i\omega\tau_1}{1 + i\omega\tau_2}. \tag{16}$$

Note that Eq. (16) has the form of a conventional relaxation[10] with τ_1 and τ_2 analogous to τ_{VS} and τ_{PS}, the isentropic relaxation times at constant volume and constant pressure, respectively. The unknowns τ_1 and τ_2 are found by inserting Eq. (16) into (5) and then matching coefficients to as high a power of ω as possible, in this case up to and including quadratic terms. The resulting relaxation times are

$$\tau_1 = \frac{(\gamma-1)(\tau_v - \tau_h)\tau_h}{\tau_v + (\gamma-1)\tau_h}, \tag{17}$$

$$\tau_2 = \tau_1 + \tau_v + (\gamma-1)\tau_h. \tag{18}$$

The relaxation strength becomes

$$\varepsilon = \frac{\tau_2 - \tau_1}{\tau_2} = \frac{[\tau_v + (\gamma-1)\tau_h]^2}{[\tau_v + (\gamma-1)\tau_h]^2 + (\gamma-1)(\tau_v - \tau_h)\tau_h}. \tag{19}$$

Because the large magnitude of the relaxation strength invalidates the usual assumption that $\alpha/k_0 \ll 1$, the exact expressions for the sound speed ratio and absorption must be used here:[11]

$$\frac{c_0}{c} = \frac{1}{\sqrt{2}}\left\{1 - \varepsilon\frac{\omega^2\tau_2^2}{1+\omega^2\tau_2^2} + \left[\left(1 - \varepsilon\frac{\omega^2\tau_2^2}{1+\omega^2\tau_2^2}\right)^2 + \left(\varepsilon\frac{\omega\tau_2}{1+\omega^2\tau_2^2}\right)^2\right]^{1/2}\right\}^{1/2}, \tag{20}$$

$$\frac{\alpha}{k_0} = \frac{1}{\sqrt{2}}\left\{-\left(1 - \varepsilon\frac{\omega^2\tau_2^2}{1+\omega^2\tau_2^2}\right) + \left[\left(1 - \varepsilon\frac{\omega^2\tau_2^2}{1+\omega^2\tau_2^2}\right)^2 + \left(\varepsilon\frac{\omega\tau_2}{1+\omega^2\tau_2^2}\right)^2\right]^{1/2}\right\}^{1/2}. \tag{21}$$

Plots of the sound speed ratio and absorption derived from the R_{11} approximant, Eqs. (20) and (21), are shown as dashed lines in Figs. 1(a) and 1(b). Approximate and exact theory agree up to *f/P* ratios of 3×10^9 for the sound speed ratio and 10^9 Hz/atm for the absorption. If the thermal conductivity is neglected ($\tau_h = 0$), then the Padé R_{11} approximant yields the exact solution. For relaxation processes with $\varepsilon \ll 1$, Eqs. (20) and (21) reduce to the well-known simpler forms.

Equations (17)–(21) reveal several features of interest. First, since in gases $\tau_v \approx \tau_h$, it follows that $|\tau_1| \ll |\tau_2|$. Second, since $\tau_v < \tau_h$ for most gases, it follows that $\tau_1 < 0$, a fact which has no bearing on Eqs. (20)–(21) since both equations are independent of τ_1. However, this causes $\varepsilon > 1$, thus yielding a much larger dispersion step (100%) and normalized absorption peak ($\sim$0.36) than found in internal molecular relaxations. Finally, the reader can easily verify that in the low-frequency limit Eqs. (18), (19), and (21) yield the Stokes–Kirchhoff formula for the classical absorption.

Equation (18) permits one to determine the translational collision number Z_{tr}, i.e., the number of collisions required for the translational velocities to advance by $1/e$ toward their equilibrium distribution following a disturbance. First it is necessary to derive the isothermal relaxation time τ_{2T}, which is the appropriate relaxation time for finding collision numbers,[12] from the isentropic relaxation time τ_2, by multiplying the latter by the ratio of low-frequency to high-frequency limits of C_P:

$$Z_{tr} = \tau_{2T}/\tau_c = \tau_2 C_{P0}/(\tau_c C_{P\infty}). \tag{22}$$

For argon at 20 °C, 1 atm, the translational relaxation time τ_2, the mean free time τ_c, and specific heat ratio are 2.96×10^{-10} s, 1.72×10^{-10} s, and (5/2)/(4/2)=1.25, respectively, yielding a collision number $Z_{tr}=2.2$. Obviously translational velocities approach equilibrium very rapidly.

The translational dispersion and absorption formulas are applicable to polyatomic gases, but here one must carefully consider the role of rotational relaxation.[13] As shown in Ref. 10, the relaxation formulas for c_0^2/c^2 and $(\alpha/k_0)(c_0^2/c^2)$ in the case of multiple relaxations are additive. If the rotational collision number is far in excess of the translational collision number, then the decoupled relaxation times can be used. This is not the case for air, for which the translational collision number at 20 °C, 1 atm is $Z_{tr}=1.6$ using Eqs. (18) and (22), but the rotational collision number is determined to be $Z_{rot}=4.8$ from a formula given by Bass and Sutherland.[14] Thus, since the collision numbers and subsequent relaxation times are too close together for the two relaxations to be considered independent, a multiple relaxation model must be used. Equations (18)–(19) provide a means of facilitating this task. For example, if the relaxations are considered to occur in parallel, then the formula for parallel relaxations[1] yields collision numbers equal to 1.5 and 5.1.

Bauer[15] approaches the coupled translational-rotational relaxation problem by introducing complex specific heats and thermal conductivity (but not viscosity) into the complex sound speed, and suggests an iterative procedure to solve for the dispersion and absorption, but he does not obtain collision numbers.

The above translational relaxation formulas are also applicable to liquids since replacement of Eq. (1) by an equation of state for liquids will still lead to Eqs. (17) and (18).

III. TRANSITION TO HIGH-FREQUENCY/PRESSURE RATIOS

Classical continuum analysis is based on the assumption that the molecular mean free path is much smaller than the acoustical enclosure containing the test gas. As a result classical kinetic theory yields transport coefficients that are independent of pressure. With decreasing pressure (increasing

frequency/pressure ratio), however, the growth of the molecular mean free path relative to the enclosure dimensions leads to a pressure dependence of the transport coefficients, whereby the molecular mean free path approaches an asymptotic limit with vanishing pressure. The one-dimensional model developed here accounts for the finite limit imposed upon the mean free path by the walls of the enclosure.

Let the transmitter be located at $x=0$ and the receiver at $x=h$. Within an infinitesimal element dx, at any instant, half the molecules are moving in the $+$x-direction and half in the $-x$-direction. Each molecule is destined to have an individual free path ℓ, i.e., distance traveled prior to a collision. However, if a free path is sufficiently long, then a molecule will collide with a wall before it collides with another molecule. Wall intervention shortens the mean free path. If the molecular free paths are exponentially distributed[16] and the molecular spatial density is uniform, then the enclosure-bounded mean free path ℓ_c' of the positive-going molecules is given by

$$\ell_c' = \frac{1}{h\ell_c}\left\{ \int_0^h \left[\int_0^{h-x} \ell \exp(-\ell/\ell_c) d\ell + \int_{h-x}^{\infty} (h-x) \times \exp(-\ell/\ell_c) d\ell \right] dx \right\}, \tag{23a}$$

$$= \ell_c \left\{ 1 - \left[\frac{1-\exp(-r)}{r} \right] \right\}, \tag{23b}$$

where ℓ_c is the unbounded mean free path, and $r=h/\ell_c$, which is proportional to pressure since ℓ_c scales inversely with pressure. A similar analysis leads to an identical result for the negative-going molecules. In the limit of infinite pressure, $\ell_c' \to \ell_c \to 0$. Likewise, in the limit of vanishing pressure, $\ell_c \to \infty$ but $\ell_c' \to h/2$. The last result is consistent with the expectation that the mean molecular position is halfway between the walls. Then, if ℓ_c' is used in place of ℓ_c in the simple kinetic-theoretical expressions for the transport coefficients,[17] one obtains the following expressions for the pressure dependence of the transport coefficients and associated relaxation times:

$$\kappa' = \kappa \left\{ 1 - \left[\frac{1-\exp(-r)}{r} \right] \right\}, \tag{24a}$$

$$\mu' = \mu \left\{ 1 - \left[\frac{1-\exp(-r)}{r} \right] \right\}, \tag{24b}$$

$$\tau_v' = \tau_v \left\{ 1 - \left[\frac{1-\exp(-r)}{r} \right] \right\}, \tag{24c}$$

$$\tau_h' = \tau_h \left\{ 1 - \left[\frac{1-\exp(-r)}{r} \right] \right\}, \tag{24d}$$

$$\tau_1' = \tau_1 \left\{ 1 - \left[\frac{1-\exp(-r)}{r} \right] \right\}, \tag{24e}$$

$$\tau_2' = \tau_2 \left\{ 1 - \left[\frac{1-\exp(-r)}{r} \right] \right\}, \tag{24f}$$

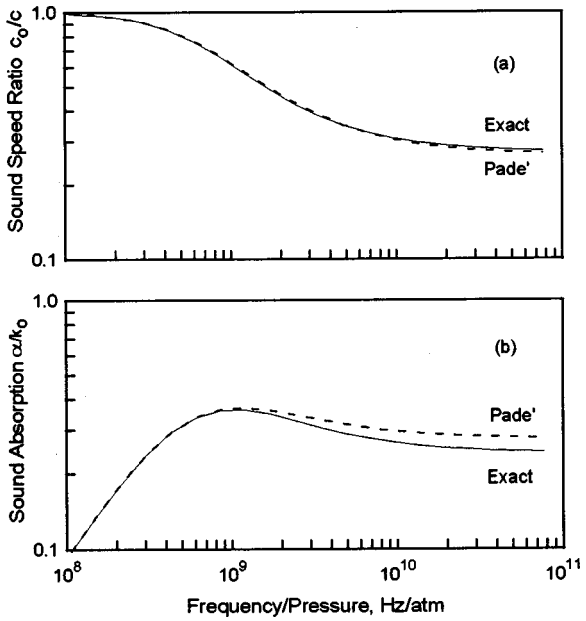


FIG. 2. Classical (a) dispersion and (b) absorption of sound in argon at 20 °C, incorporating the pressure dependence of the transport properties, Eqs. (24a)–(24f), versus frequency/pressure ratio. Solid lines: exact theory, Eqs. (14) and (15). Dashed lines: approximate (Padé) theory, Eqs. (20) and (21).

in which primed quantities incorporate the above-derived pressure dependence. The relaxation strength remains independent of pressure, as may be seen from Eq. (19), because both decoupled relaxation times have the same pressure dependence. By replacing the unprimed relaxation times, Eqs. (6), (7), and (18) with the primed relaxation times Eqs. (24c), (24d), and (24f) in Eqs. (8)–(15) and (20)–(21), one accounts for the pressure dependence of the transport properties in the sound speed and absorption. The resulting plots are shown in Fig. 2(a) and (b). The approximate and exact sound speed ratios are practically indistinguishable, approaching asymptotic limits of 0.294 and 0.298, respectively. The approximate normalized absorption is slightly greater than the exact, approaching 0.290 versus 0.259. These trends reflect the fact that the relaxation time now approaches a finite limit instead of increasing indefinitely with vanishing pressure.

IV. COMPARISON TO EXPERIMENT

Figures 3(a) and 3(b) show plots of the theoretical sound speed ratio and absorption in argon, based on the Padé R_{11} approximation [Eqs. (20), (21), and (24)], together with selected data.[18–20] No adjustments have been made to fit the data; rather, all argon parameters entering Eqs. (6) and (7) are handbook values at 20 °C, 1 atm. The enclosure dimension h, a key parameter, is obtained directly or inferred from the authors' data. For the data of Refs. 18, 19, and 20 the values of ωh are taken to be 2488, 3000, and 1970 m/s, leading to values of $h=3.6\times10^{-5}$, 0.004 775, and 0.000 322 m, respectively. Since the values of ωh are remarkably consistent, in view of the wide range of ω, a mean value of $\omega h=2500$ m/s was used for the plots of Fig. 3.

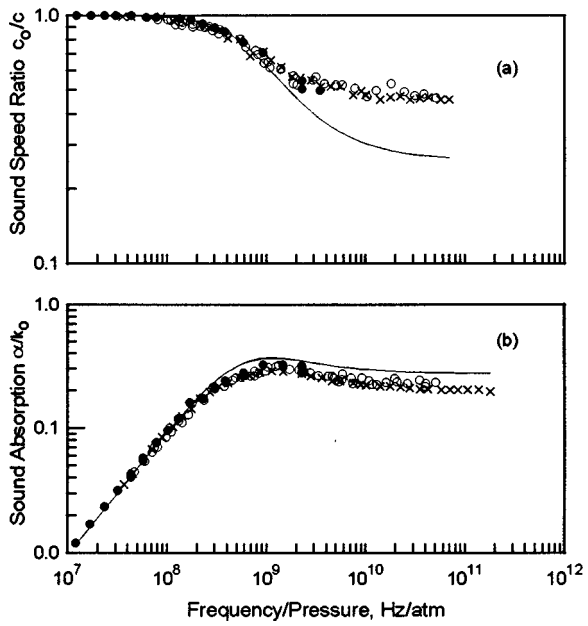


FIG. 3. Experimental and theoretical (a) dispersion and (b) absorption of sound in argon versus frequency/pressure ratio. Solid line: approximate (Padé) theory, Eqs. (20) and (21), incorporating the pressure dependence of the transport properties, Eqs. (24a)–(24f). Symbols: experimental—solid circles, Ref. 18; open circles, Ref. 19; crosses, Ref. 20. Temperatures are 20 °C for the theory, 20–23 °C for the experiments.

The theoretical sound speed ratio and absorption show good agreement with experiment from the classical through the transition region (i.e., up to $f/P \approx 10^9$ Hz/atm). The theoretical sound speed ratio lies about a factor of 1.7 below the data in the collisionless region and the theoretical absorption about a factor of 1.2 above the data. Adjustments in h will permit either the sound speed or absorption to fit the data, but not both simultaneously. Thus it is concluded that the data lying beyond the transition region cannot be represented as a single relaxation process.

These results are consistent with the model described by Schotter.[20] At small molecular mean free paths, relative to the transmitter–receiver distance, the propagating modes are the familiar ''collective'' modes established primarily by binary collisions. As the mean free path grows, the collective modes give way to single-particle ''quasi-modes'' characterized by collisionless propagation between the transmitter and receiver. The transition from collective mode to single-particle mode propagation occurs at transmitter–receiver distances of one mean free path for the dispersion but one-half mean free path for the absorption. From Schotter's Fig. 6, one can infer that these transitions take place at frequency/pressure ratios in the vicinity of 7.0×10^9 and 1.4×10^{10} Hz/atm, respectively. Because the theory developed here assumes collective-mode propagation over the full range of frequency/pressure ratios, it does not account for the level of dispersion or absorption associated with single-particle mode propagation.

V. CONCLUSIONS

The phenomenological theory of the translational relaxation times in gases, based on a solution to the secular equation by means of the Padé R_{11} approximant, illustrates that translational relaxation can be formulated in terms of a conventional relaxation process, whereby well-defined relaxation times evolve naturally from the phenomenological equations. Even at this low order of approximation the solution is in remarkable agreement with the exact solution, at least over the classical and transition regions of *f/P*, and yields the familiar dispersion step and absorption peak. This formulation facilitates the analysis in cases where translational relaxation is but one component of a multiple relaxation process, a simple case being translational-rotational coupling. It permits the determination of translational collision numbers and is not beset by the convergence difficulties inherent in power series solutions to the Boltzmann transport equation.

At sufficiently high values of *f/P* the experimental sound speed ratio and absorption both approach limits which cannot be explained by classical continuum theory. The physical argument employed here, namely that these limits are due to the molecular mean free paths being bounded by the enclosure dimensions, is supported by experiment only through the transition region and not over the full range of *f/P*. The discrepancy at higher *f/P* ratios is attributable to the difference between the type of propagation mode prevailing in the experiments (collisionless, confined) and that assumed by the theory (collective, free-field).

ACKNOWLEDGMENTS

Comments by my Langley colleagues, Willard E. Meador and Fereidoun Farrasat, are gratefully acknowledged.

[1] A. B. Bhatia, *Ultrasonic Absorption* (Dover, New York, 1967).
[2] M. Greenspan, ''Transmission of sound waves in gases at very low pressures,'' in *Physical Acoustics*, Vol. IIA, edited by W. P. Mason (Academic, New York, 1965), pp. 1–45.
[3] F. V. Hunt, ''Propagation of sound in fluids,'' in *American Institute of Physics Handbook*, 3rd ed. (McGraw-Hill, New York, 1972), pp. 3-37 to 3-68.
[4] C. Truesdell, ''Precise theory of the absorption and dispersion of forced plane infinitesimal waves according to the Navier–Stokes equation,'' J. Rational Mech. Anal. **4**, 643–741 (1953).
[5] J. J. Markham, R. T. Beyer, and R. B. Lindsay, ''Absorption of sound in fluids, '' Rev. Mod. Phys. **23**, 353–411 (1951).
[6] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968), Chap. 6.
[7] A. D. Pierce, *Acoustics: An Introduction to its Physical Principles and Applications* (Acoustical Society of America through the American Institute of Physics, Woodbury, NY, 1981), p. 52.
[8] J. P. M. Trusler, *Physical Acoustics and Metrology of Fluids* (Adam Hilger: IOP Publishing Ltd., Bristol and Philadelphia, 1991), Chap. 2.
[9] R. F. Churchhouse, Ed., *Handbook of Applicable Mathematics* (Wiley, New York, 1981), pp. 216–223.
[10] H. J. Bauer, ''Phenomenological theory of the relaxation phenomena in gases,'' in *Physical Acoustics*, Vol. IIA, edited by W. P. Mason (Academic, New York, 1965), pp. 47–131, see Eq. (58).
[11] K. F. Herzfeld and T. A. Litovitz, *Absorption and Dispersion of Ultrasonic Waves* (Academic, New York, 1959).
[12] H. O. Kneser, *Physical Acoustics*, Vol. IIA, edited by W. P. Mason (Academic, New York, 1965), pp. 142–144.
[13] M. Greenspan, ''Rotational relaxation in nitrogen, oxygen, and air,'' J. Acoust. Soc. Am. **31**, 155–160 (1959).
[14] H. E. Bass and L. C. Sutherland, ''On the rotational collision number for air at elevated temperatures,'' J. Acoust. Soc. Am. **59**, 1317–1318 (1976).
[15] H. J. Bauer, ''Influences of transport mechanisms on sound propagation in gases,'' Adv. Mol. Relax. Processes **2**, 251–318 (1972).
[16] E. Nasser, *Fundamentals of Gaseous Ionization and Plasma Electronics*

(Wiley-Interscience, New York, 1971); J. D. Cobine, *Gaseous Conductors, Theory and Engineering Applications* (Dover, New York, 1958).
[17] J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids* (Wiley, New York, 1954).
[18] M. Greenspan, ''Propagation of sound in five monatomic gases,'' J. Acoust. Soc. Am. **28**, 644–648 (1956). An average enclosure dimension $h=36$ μm is assumed for the present analysis.
[19] E. Meyer and G. Sessler, ''Schallausbreitung in Gasen bei hohen Frequenzen und sehr niedrigen Drucken,'' Z. Phys. **149**, 15–39 (1957).
[20] R. Schotter, ''Rarefied gas acoustics in the noble gases,'' Phys. Fluids **17**, 1163–1168 (1974).

Resonant oscillation of a liquid metal column driven by electromagnetic Lorentz force sources

Sergey Makarov, Reinhold Ludwig,[a] and Diran Apelian
Metal Processing Institute, Department of Mechanical Engineering, Worcester Polytechnic Institute, Worcester, Massachusetts 01609-2280

(Received 21 July 1998; revised 24 November 1998; accepted 11 January 1999)

In this paper, a theoretical study is conducted in order to establish the feasibility of a liquid metal acoustic resonator (liquid gallium or liquid aluminum) for high-amplitude acoustic oscillations. The fundamental resonant frequency typically lies between 5 and 40 kHz. The oscillations are induced by an alternating Lorentz force density applied directly to the liquid metal volume. Depending on the boundary conditions, two different resonator types (open–closed and open–open) are theoretically investigated. The analysis incorporates the effects of impedance termination, volume absorption, wall friction, acoustic radiation from the open end, and nonlinear inflow–outflow losses. The actual elasticity of the container, either a ceramic or quartz tube, and the coupled solid–liquid interactions are taken into consideration. Based on this investigation, theoretical predictions are conducted for the quality factor and the pressure level for the liquid metal resonator under various geometric and boundary conditions. They indicate that resonant amplitudes of 10–20 atm can be achieved using commercially available high-current audio amplifiers. © *1999 Acoustical Society of America.* [S0001-4966(99)03104-5]

PACS numbers: 43.35.Bf, 43.35.Zc, 43.20.Ks, 43.25.Gf, 43.38.Ar, 43.38.Dv [HEB]

LIST OF SYMBOLS

a	tube radius [m]
$\mathbf{B}$, B	magnetic flux density [Wb/m^2=Tesla]
$B/A=n-1$	nonlinearity parameter of the liquid
c_0	sound speed in the liquid metal [m/s]
c_l	longitudinal wave speed in the termination/tube material [m/s]
c_p	phase velocity of compressional waves in a elastic plate [m/s]
c_p	specific heat at constant pressure [J/kg K]
E	Young's modulus of the termination/tube material [GPa]
$\mathbf{f}$, f, $\hat{f}$	Lorentz force density and its complex amplitude [N/m^3]
f_n	driving force density for nth resonance [N/m^3]
F	total Lorentz force [N]
F_n	total driving force for nth resonance [N]
h	wall thickness of the container [m]
$\mathbf{i}, i_x$	current density [A/m^2]
I, $\hat{I}$	total current (not rms!) and its complex amplitude [A]
j	imaginary unit
k_1, k_{1n}	wave number of the driving force [1/m]
K	circumferential stiffness of the tube wall [N/m^3]
L	height of the liquid metal column [m]
p, p_0	acoustic pressure and ambient pressure, respectively [Pa]
P	pressure amplitude of resonant oscillations (at antinode) [atm, Pa]
Q	quality factor of the resonator
$\mathbf{r}, \mathbf{r}_0$	observation point and source location point, respectively
S	tube cross-section [m^2]
w	velocity of the liquid metal in z direction [m/s]
x,y,z	Cartesian coordinates [m]
$Z_s=\rho_s c_l$	termination acoustic impedance [kg/m^2 s]
α	semi-empirical loss factor for inflow–outflow losses
α/f^2	sound absorption coefficient [s^2/m]
β	thermal expansion coefficient at constant pressure [1/K]
η	radial displacement of the tube wall [m]
η_s	shear viscosity of the liquid metal [Pa s]
η_e	effective shear viscosity of the liquid metal [cf. Eq. (1)] [Pa s]
θ	polar angle (xy-plane)
κ	thermal conductivity of the liquid metal [J/(m s K)]
λ	acoustic wavelength [m]
μ_e	effective kinematic viscosity (η_e/ρ_0) of the liquid metal [m^2/s]
ν	Poisson's coefficient of the termination/tube material
ρ_0	density of the liquid metal [kg/m^3]
ρ_s	density of the termination/tube material [kg/m^3]
ρ_l	resistivity of the liquid metal [Ω m]
ω	angular frequency [1/s]

INTRODUCTION

Acoustic resonators are valuable tools for collecting small inclusions in various fluids at nodes (anti-nodes) due to

[a]Also at Electrical and Computer Engineering Department, Worcester Polytechnic Institute, Worcester, MA 01609-2280.

the exerted acoustic radiation force.[1–3] The intense acoustic field can increase the wettability of solid inclusions in a liquid metal[4] and thus facilitate their deposition on the sensor surface. This makes acoustic standing waves an attractive mechanism from the viewpoint of sensor development for processing small inclusions in molten metals. Detection of small nonmetallic inclusions residing in a metallic melt (aluminum) has received increased attention due to industrial demand for the manufacturing of high quality metal products.[5–7]

To induce high-amplitude resonant oscillations of a liquid metal column, two problems have to be investigated. The first is related to the excitation force. Standard piezoelectric excitation becomes cumbersome under high-temperature conditions. Guided buffer rods and special coolers are necessary to transmit the pressure wave to the melt.[8–10] Alternatively, excitation methods based on electromechanical resonance show good performance, but they are restricted to one fixed frequency.[1]

The second problem pertains to the high acoustic impedance of the liquid metal, typically 6–9 times greater than the impedance of water. If a buffer rod is used as a transmitting element, its acoustic impedance is within the same order of magnitude. Therefore, no effective reflection from the driven end is created. This condition considerably limits the oscillation amplitude of a piezoelectric resonator. In practice, it is difficult to achieve good acoustic reflection from *any* rigid termination in molten metal if such a termination is applied to the end of the tube. Moreover, liquid metal also stretches the confining tube boundaries, making it necessary to take into account the local wall response.

The focus of the present paper is a theoretical study of the resonance conditions for a liquid metal column in a small tube. The emphasis is geared toward the two special problems as outlined above. We propose an alternative driving mechanism based on a periodic Lorentz force induced by an external static magnetic field. This force directly acts on the liquid metal in the inner volume of the resonator. Driving frequencies range from several kHz to several tens of kHz. Two resonator types are basically investigated: a quarter wavelength open–closed resonator with rigid termination at one end, and a half-wavelength open–open resonator. Our goal is to calculate the quality factor and the resonant frequency under different operating conditions and then propose the better choice for the resonator construction.

Since the wavelength is always large in comparison with the tube radius, a spatially one-dimensional acoustic model is formulated, which includes different absorption mechanisms. For the second resonator type, a nonlinear model of the inflow–outflow loss becomes necessary to correctly predict the oscillation amplitude. The force distribution analysis involves analytical calculations and simple eigenfunction expansions.

The paper is organized as follows. Section I outlines the problem statement and gives an overview of the material properties. Section II deals with practically achievable spatial force distributions and the corresponding eigenfunction expansion. Section III introduces the linear acoustic theory. A nonlinear extension is considered in Sec. IV and subsequently applied to the second resonator type. Finally, discussion and conclusions are presented in Sec. V.

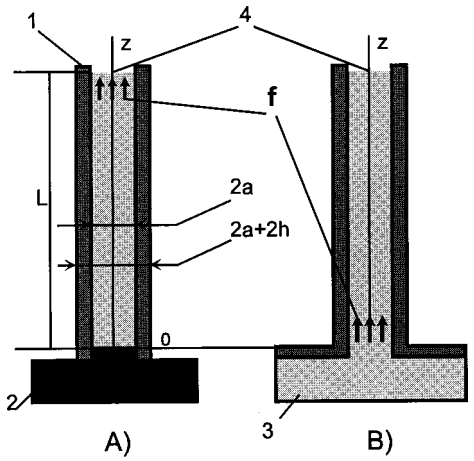


FIG. 1. Geometry of two resonator types. 1—ceramic tube; 2—rigid foundation for the open–closed resonator; 3—outer liquid volume for the open–open resonator; 4—free surface of liquid metal.

I. PROBLEM STATEMENT

The resonator geometry is shown schematically in Fig. 1. Liquid metal fills a cylindrical tube (1) to a column of height L. Tube (1) with inner radius a and outer radius $a+h$ is either closed through a rigid massive foundation (2) or it is connected with an outer liquid volume (3). The upper surface of the liquid column (4) always satisfies the free surface condition by being subjected to an ambient static pressure p_0. The construction shown in Fig. 1 on the left represents a quarter wavelength open–closed resonator (type A) when the melt is driven by a periodic force at its eigenfrequencies. The construction shown in Fig. 1 on the right is a half-wavelength open–open resonator (type B). We will study both types and analyze their performance in terms of quality factor and maximum pressure amplitude.

As a driving force, a distributed Lorentz force density **f** will be analyzed, concentrated mainly in the vicinity of the pressure nodes (Fig. 1) for both resonator types. Only fundamental and second resonance modes are the subject of this study. This approximately yields $L=\lambda/4,3\lambda/4$ for type A, and $L=\lambda/2,\lambda$ for type B. The typical scales under study are $L=50-300\,\mathrm{mm}$, $a=2.5-10\,\mathrm{mm}$, and $h=1-2.5\,\mathrm{mm}$. The liquid metals investigated in this paper are liquid gallium and liquid aluminum.

Since $\lambda \gg a$ is in all situations, a spatially one-dimensional model of resonant acoustic vibration is implemented. The one-dimensional model straightforwardly includes the effect of the finite termination impedance at the tube ends, the effect of thermoviscous sound absorption in the inner tube volume, and the hydrodynamic nonlinearity effect. After some averaging manipulations it is also possible to take into account the effect of wall friction and the effect of locally reacting elastic tube walls.

To get a preliminary estimate of the relative importance of those effects for the acoustic resonator performance, we have collected acoustic and material properties for liquid gallium and liquid aluminum in Tables I and II.[11–17] Necessary

TABLE I. Mechanical and thermal properties of liquid gallium and liquid aluminum near the melting point. For the purpose of comparison, data for water at 20 °C are also presented. The columns denote, respectively, melting point MP in °C of the liquid, density ρ_0, specific heat at constant pressure c_p, ratio of specific heats $\gamma=c_p/c_v$, sound speed c_0, shear viscosity η_s, thermal conductivity κ, sound absorption coefficient α/f^2, effective shear viscosity η_e, and nonlinear parameter of liquid B/A.

	MP, °C	ρ_0, kg/m^3	c_p, J/kg K	c_p/c_v	c_0, m/s	η_s, Pa s	κ, J/m s K	α/f^2, s^2/m	η_e/η_s	B/A
Ga	29.78	6100	400	1.10	2800	0.0021	31	1.6$e-$15	2.5	3.7
Al	660.2	2400	1100	1.25	4700	0.002	115	2.3$e-$15	8.9	2.1
H_2O	0	1000	4200	1.0	1480	0.001	0.5	25.0$e-$15	1.0	5

data on tube and termination materials are given in Table III.[17–19] Among other familiar quantities, the effective shear viscosity

$$\eta_e=\eta_s\left(1+\frac{\beta c_0^2}{c_p}\sqrt{\frac{\kappa}{\eta_s c_p}}\right)^2, \tag{1}$$

is introduced into consideration to study the wall friction effect[20] in the thermoviscous fluid [Ref. 21, p. 209; Ref. 22, Eq. (36)]. This characterizes equivalent increase in shear viscosity while taking into account the temperature boundary layer (cf. Ref. 22). In comparison with a water- or air-filled resonator, the following observations can be made from Tables I–III:

(i) Due to the extremely high acoustic impedance of the liquid metal it is difficult to guarantee high reflectivity from the closed end boundary. The most suitable material for this purpose with largest longitudinal impedance would be probably tungsten or tungsten carbide.

(ii) The same reason makes it necessary to include into consideration the local stiffness and mass response of the tube wall.

(iii) The effect of thermoviscous volume absorption seems to be unimportant in comparison with the acoustic losses due to wall friction.

II. DRIVING FORCE

A. Geometry

A typical situation allows us to consider two current supplying electrodes immersed in the liquid metal in proximity to the tube wall, as is shown in Fig. 2. For simplicity, we have put the origin of a Cartesian coordinate system in the plane of the two electrodes. The external static magnetic field **B** is directed toward the *y*-axis. The Lorentz force density is given by

$$\mathbf{f}=\mathbf{i}\times\mathbf{B}. \tag{2}$$

TABLE II. Physical properties of liquid gallium and liquid aluminum near the melting point. The columns denote, respectively, surface tension of liquid state σ with respect to water (σ_0 at 20 °C), wetting properties, and electric resistivity ρ_l in Ω m.

Metal	σ/σ_0	Wetting	ρ_l, Ω m
Ga	9.7	glass, porcelain	25.9×10^{-8}
Al	13	quartz	26.0×10^{-8}

For our consideration, only the *z*-component of **f**, denoted by *f*, is of interest for the excitation of longitudinal vibrations. From Eq. (2), we obtain

$$f=Bi_x. \tag{3}$$

Quasi-steady[23] current distribution within a finite conducting cylinder is calculated based on the Green's function approach.[24,25] Green's function of a simple current source $q=I\delta(\mathbf{r}-\mathbf{r}_0)$ in the interior of the finite cylinder of height *L* is obtained in terms of cylindrical coordinates r,θ,z in the form (cf. also the corresponding acoustic solutions[26,27])

$$G(\mathbf{r}|\mathbf{r}_0)=I\rho_l\sum_{m=0,n=0,p=0}^{\infty}G_{mnp}H_{mnp}(\mathbf{r})H_{mnp}(\mathbf{r}_0), \tag{4}$$

$$H_{mnp}(\mathbf{r})=J_m(\lambda_{mn}r/a)\cos(m\theta)\cos(\pi pz/L), \tag{5}$$

$$G_{mnp}=\frac{\epsilon_m\epsilon_p}{\pi La^2(1-m^2/\lambda_{mn}^2)(\lambda_{mn}^2/a^2+\pi^2p^2/L^2)J_m^2(\lambda_{mn})}. \tag{6}$$

Here, J_m is the Bessel function of order *m*, λ_{mn} are consecutive zeros of the Bessel function derivative J_m', i.e., $J_m'(\lambda_{mn})=0$, $n=0,1,\ldots$. The Neumann factors ϵ_m, ϵ_p are equal to *l* if *m*, $p=0$ and 2 otherwise. Figure 3 shows the contour plot of the axial current density component i_x, which is proportional to the Lorentz force density, due to two current sources oppositely placed on the wall, along the symmetry plane. The current and the force are well localized near the plane of two electrodes. The width of the localization domain is about one tube radius *a*.

The total Lorentz force *F* in the *z* direction is of primary importance for the oscillation magnitude. It follows from the current conservation law that

$$F=B\int_V i_x\,dx\,dy\,dz=2aIB. \tag{7}$$

B. Eigenfunction expansion

In general, we assume a harmonic current and harmonic driving force, i.e., $I=\hat{I}e^{j\omega t}$, $f=\hat{f}e^{j\omega t}$. A connection should be established between the physically available spatial force distribution and the eigenfunction expansion, used for the theoretical predictions below. The eigenfunction expansion is governed by

TABLE III. Mechanical properties of a range of solid insulators for tube walls and/or tube terminations. For the purpose of comparison, data of the acoustic impedances of liquid gallium and aluminum are also presented. The columns denote, respectively, density ρ_s, Young's modulus E, Poisson's ratio ν, low-frequency phase velocity of compressional waves in a elastic plate c_p, longitudinal impedance $Z_s=\rho_s c_l$, and maximal use temperature, respectively.

Material	ρ_s, kg/m^3	E, Gpa	Poisson's ratio	c_p, m/s	$Z_s=\rho_s c_l$, kg/m^{2}s	Maximal operating temperature, °C
Tungsten carbide (WC)	15 700	700	0.18	6788	1.1e+8	>1000 °C
Tungsten	19 350	375	0.3	4630	1.1e+8	>1000 °C
Friatec-Degussit® alumina ceramics AL-23	3800	380	0.22	10 300	4.1+7	1950
Fused quartz	2200	73	0.17	5850	1.3e+7	1000
Liquid Ga	6100	⋯	⋯	⋯	1.7e+7	⋯
Liquid Al	2400	⋯	⋯	⋯	1.1e+7	⋯

$$f=e^{j\omega t}\sum_{n=1}^{\infty} f_n \sin k_{1n}z, \quad k_{1n}=\left(n-\frac{1}{2}\right)\frac{\pi}{L}, \quad z\in(0,L), \tag{8a}$$

$$f=e^{j\omega t}\sum_{n=1}^{\infty} f_n \cos k_{1n}z, \quad k_{1n}=n\frac{\pi}{L}, \quad z\in(0,L) \tag{8b}$$

for the open–closed or open–open resonator, respectively.

Since $a\ll L$, the axial Lorentz force density, Eq. (3), for the lowest resonances is quite well approximated by a δ-function located in the plane of two electrodes at $z=z_0$, i.e.,

$$\hat{f}=\frac{F}{S}\delta(z-z_0)=\frac{2a\hat{I}B}{S}\delta(z-z_0). \tag{9}$$

We will put $z_0=L-a$ for the open–closed resonator and $z_0=a$ for the open–open resonator, respectively. This gives in the main order with respect to small parameter a/L

$$\begin{aligned} f_n&=\frac{4a\hat{I}B}{SL}(-1)^{n+1} \quad \text{for type A,}\\ f_n&=\frac{4a\hat{I}B}{SL} \quad \text{for type B.}\end{aligned} \tag{10}$$

Due to the orthogonality condition for the corresponding boundary value problem, coefficient f_1 characterizes the fundamental resonance only, coefficient f_2 the second resonance, etc. We restrict ourselves to the fundamental or the second resonance expansion term.[28]

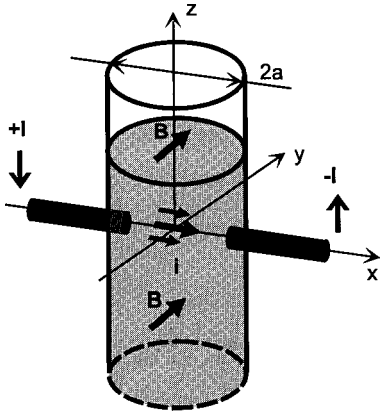


FIG. 2. Two current carrying electrodes and the external magnetic field applied to the liquid metal column.

III. LINEAR ACOUSTIC THEORY

A. Wave equation for the liquid metal

The wave equation for the liquid metal should thus involve the local wall response and the boundary layer effect. If the acoustic pressure stretches the flexible tube so that its radius a becomes $a+\eta$ ($\eta/a\ll 1$), the wave equation is simply obtained from mass and momentum conservation laws in the form (Ref. 29, pp. 688–689)

$$\frac{\partial^2 p}{\partial t^2}-c_0^2\frac{\partial^2 p}{\partial z^2}+c_0^2\frac{2\rho_0}{a}\frac{\partial^2 \eta}{\partial t^2}=0. \tag{11}$$

In Eq. (11), we neglected nonlinearities such as $p\eta$.

The effect of wall friction (long-wavelength limit) results in the substitution[20]

$$\frac{\partial p}{\partial z}\to\left[1-\frac{2}{a}\sqrt{\frac{\mu_e}{j\omega}}\right]\frac{\partial p}{\partial z}, \tag{12}$$

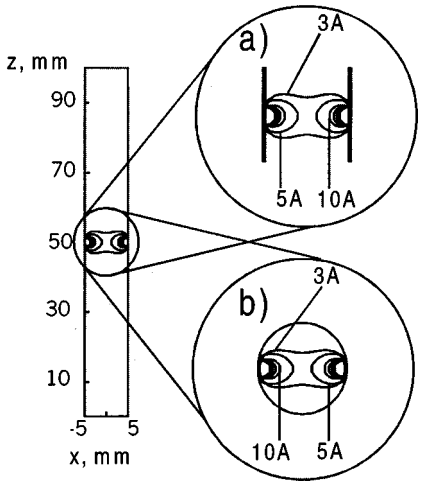


FIG. 3. Contour plot of axial current density i_x for the cylinder of 100 mm in height and 10 mm in width. (a) Axial plane; (b) radial plane. Two point current sources are oppositely placed on the tube wall. The total current is 35 A rms.

which should be made in the momentum equation for a harmonic wave. The continuity equation remains the same. In time domain, such substitution corresponds to the integral operator[30]

$$\frac{\partial p}{\partial z} \rightarrow \frac{\partial p}{\partial z} - \frac{2}{a}\sqrt{\frac{\mu_e}{\pi}}\int_{-\infty}^{t} \frac{\partial p(z,\tau)/\partial z}{\sqrt{t-\tau}}\, d\tau. \tag{13}$$

Equation (11) is then derived in the form:

$$\frac{\partial^2 p}{\partial t^2} - c_0^2\left(\frac{\partial^2 p}{\partial z^2} - \frac{2}{a}\sqrt{\frac{\mu_e}{\pi}}\int_{-\infty}^{t} \frac{\partial^2 p(z,\tau)/\partial z^2}{\sqrt{t-\tau}}\, d\tau\right) + c_0^2 \frac{2\rho_0}{a}\frac{\partial^2 \eta}{\partial t^2} = 0. \tag{14}$$

B. Vibrations of tube walls

The local mass response of the cylindrical tube with wall thickness h is governed by the equation[29,31,32]

$$\frac{\partial^2 \eta}{\partial t^2} - \frac{p}{\rho_s h} = 0, \tag{15}$$

which is a direct consequence of Newton's second law for the tube material under pressure load. The local stiffness response is obtained from Hook's law

$$\rho_s h \frac{\partial^2 \eta}{\partial t^2} = -K\eta. \tag{16}$$

The stiffness of the tube wall K explicitly in terms of tube parameters reads[20,32]

$$K = \frac{h}{a^2}E \text{—nontethered tube}, \qquad K = \frac{h}{a^2}\frac{E}{1-\nu^2} \text{—tethered tube}. \tag{17}$$

We are basically interested here in the nontethered tube, which is mostly free with respect to longitudinal displacements (cf. Fig. 1). Combination of Eqs. (15) and (16) yields

$$\frac{\partial^2 \eta}{\partial t^2} + \frac{K}{\rho_s h}\eta - \frac{p}{\rho_s h} = 0. \tag{18}$$

The mass response is usually of negligibly small contribution in comparison with the stiffness response.[20] However, this can be important in our case for not very hard materials with smaller Young's modulus (fused quartz) at higher frequencies of order 20 kHz.

The bending stress would give a correction

$$c_p^2 \frac{h^2}{12}\frac{\partial^4 \eta}{\partial z^4} \tag{19}$$

into the left-hand side of Eq. (18). This result is obtained from the complete equations for thin shells.[32,33] As compared to other terms in Eq. (18), such a correction is, however, of order h^2/λ^2. In our situation $h^2/\lambda^2 \ll 1$, and the contribution given by Eq. (19) can be neglected.

Equation (18) goes along with the more precise theory of acoustic wave propagation in fluid-filled elastic tubes[34–36] in the low-frequency limit and is originally due to Morse and Ingard (Ref. 29, pp. 688–689). Its more complicated versions with taking into account bending effects are also available from the literature.[33,37]

For the nontethered tube, longitudinal stresses and velocities of the wall material can be shown to be in phase with fluid pressures and velocities.[20] Therefore, they will automatically satisfy the same boundary conditions as those for the liquid metal column. This means no longitudinal stress at the open end and vanishing longitudinal displacement at the closed end (resonator type A). For the combination of a ceramic tube and tungsten termination, this is a rather good approximation to reality.

C. Coupled equations

The coupled equations for the standing wave in the flexible tube, which take into account the acoustic boundary layer effect, are obtained from the two previous subsections in the form

$$\frac{\partial^2 p}{\partial t^2} - c_0^2\left(\frac{\partial^2 p}{\partial z^2} - \frac{2}{a}\sqrt{\frac{\mu_e}{\pi}}\int_{-\infty}^{r} \frac{\partial^2 p(z,\tau)/\partial z^2}{\sqrt{t-\tau}}\, d\tau\right) + c_0^2 \frac{2\rho_0}{a}\frac{\partial^2 \eta}{\partial t^2} = -c_0^2 \frac{\partial f}{\partial z}, \tag{20}$$

$$\frac{\partial^2 \eta}{\partial t^2} + \frac{K}{\rho_s h}\eta - \frac{p}{\rho_s h} = 0. \tag{21}$$

We substitute the expression for the Lorentz force density f from Eqs. (8) of Sec. II, where only the first or the second term on the right-hand sides is retained. This corresponds to the fundamental or second resonant mode, respectively. The boundary conditions for the resonator of type A imply that

$$p(0,t)/w(0,t) = Z_s \quad \text{at } z = 0, \tag{22a}$$

$$p(L,t) = 0 \quad \text{at } z = L. \tag{22b}$$

The acoustic termination impedance Z_s is given for some hard materials in Table III. The boundary conditions for resonator type B in time domain read [cf. Ref. 21, Eq. (9.10)]

$$p(0,t) = -a\rho_0\left(\frac{8}{3\pi}\frac{\partial w(0,t)}{\partial t} - \frac{a}{2c_0}\frac{\partial^2 w(0,t)}{\partial t^2}\right) \quad \text{at } z = 0, \tag{23a}$$

$$p(L,t) = 0 \quad \text{at } z = L. \tag{23b}$$

The expression in Eq. (23a) is the linear radiation condition for the flanged tube open from the bottom into the outer liquid metal volume [cf. Fig. 1(B)].

Two quantities are to be evaluated for the fundamental and second resonance condition: the quality factor of the resonator Q and its resonant frequency Ω. The quality factor is introduced as the ratio of maximum acoustic pressure force SP on the liquid volume divided by the total driving force F_n corresponding to nth resonance (in our case $n = 1,2$). The total driving force is obtained by integration of Eqs. (8) over the entire volume if only one term is kept on the right-hand sides. As the integrand, we consider the absolute value of this term. By taking into account Eqs. (10), this gives $F_n = (2/\pi)LS|f_n| = (8/\pi)a|\hat{I}|B$ $(n = 1,2)$ and

$$Q=\frac{SP}{|F_n|}=\frac{\pi S}{8a|\hat{I}|B}P. \tag{24}$$

D. Solution: Open–closed resonator

The general solution of Eqs. (20) and (21) supplied with the corresponding boundary conditions Eqs. (22) takes the form

$$p=e^{j\omega t}(p_1\cos k_1 z+p_2\sin k_2(z-L)),$$

$$w=e^{j\omega t}\frac{1}{j\omega\rho_0}(k_1p_1\sin k_1z-k_2p_2\cos k_2(z-L)), \tag{25}$$

$$\eta=e^{j\omega t}(\eta_1\cos k_1z+\eta_2\sin k_2(z-L)).$$

Wave number k_1 is always due to the external force, i.e. [see Eq. (8b)], $k_1=\pi/2L,\ 3\pi/2L$ for the fundamental and second resonance, respectively. The free surface boundary condition is satisfied automatically. Five unknown coefficients $p_1, p_2, \eta_1, \eta_2, k_2$ are then found from the differential equations and from the remaining boundary condition. However, we will obtain the solution for each physical mechanism, limiting the resonant amplitude, separately, and then build up a combination of them. Such an approach is widely used in the study of acoustic resonant oscillations with multiple dissipation mechanisms. Furthermore, each physical effect will only be analyzed to the first order of approximation with respect to the corresponding small parameter (frequency shift or energy loss per period).

The effect of flexible tube walls is obtained if we neglect the friction term in Eq. (20) and let $Z_s\to\infty$ in Eq. (22a). This yields $p_2=0,\ \eta_2=0$ in Eq. (25). Simple calculation gives

$$\Omega=c_0k_1\left(1-\frac{\rho_0c_0^2}{a(K-\rho_shc_0^2k_1^2)+2\rho_0c_0^2}\right),\quad Q=\infty \tag{26}$$

at relatively small values of $(\Omega-c_0k_1)/c_0k_1$. Equation (26) states that there is practically no acoustic loss due to the vibrating walls. This seems to be a reasonable assumption for the liquid metal resonator surrounded by air. The effect of the acoustic boundary layer also requires $p_2=0,\ \eta_2=0$. It gives

$$\Omega=c_0k_1\left(1-\sqrt{\frac{\mu_e}{2a^2c_0k_1}}\right),\quad Q=\frac{\pi}{2L}\sqrt{\frac{a^2c_0}{2k_1\mu_e}} \tag{27}$$

to first order of approximation. Finally, the effect of nonrigid termination with zero reactance yields

$$\Omega=c_0k_1,\quad Q=\frac{\pi}{4}\frac{Z_s}{\rho_0c_0}. \tag{28}$$

The resulting resonant frequency is approximately obtained as a sum of those corrections to the unperturbed value c_0k_1. The total quality factor is well approximated by the inverse of the sum of inverse quality factors, i.e.,[38]

$$\Omega=c_0k_1+\sum_m(\Omega_m-c_0k_1),\quad Q=\frac{1}{\Sigma_mQ_m^{-1}}. \tag{29}$$

The analysis of physical conditions presented in Tables I and III indicates that the main mechanism, responsible for damping the resonant oscillations, will be the relatively low termination impedance, even if hard materials like tungsten or tungsten carbide are used. To illuminate this conclusion, Fig. 4 presents the quality factors given by Eqs. (27) and (28), respectively, as a function of height L of the liquid metal column. The fundamental resonance is studied for a tube filled with liquid gallium, assuming a tungsten or tungsten carbide termination. Nearly the same results are obtained for the tube filled with molten aluminum.

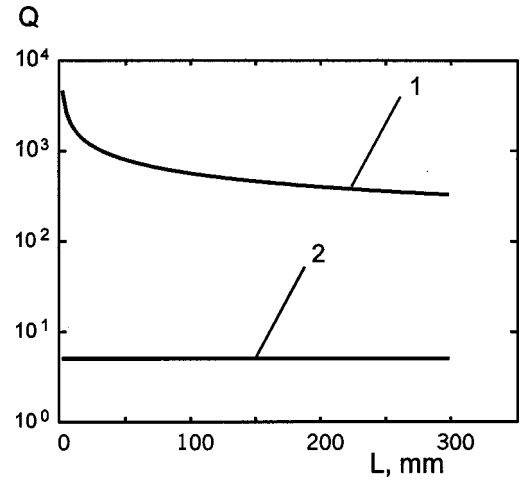


FIG. 4. Quality factor of the open–closed resonator (liquid gallium) as a function of height L of the liquid metal column at $a=5$ mm. 1—effect of the boundary layer losses; 2—effect of the finite termination impedance (tungsten or tungsten carbide).

To get some absolute estimates of achievable pressure amplitudes, a driving source has to be specified. Let us consider a commercially available ac high-current sinusoidal generator with the current amplitude of 50 A (35-A rms). If this current flows through the liquid metal subjected to an external static magnetic field about 1 Tesla (cf. Fig. 2) then the amplitude of the Lorentz force Eq. (7) appears of 0.5 N ($a=5$ mm).[39] (The magnetic-acoustic interaction[39] is assumed to be negligibly small.) The total driving force F_n corresponding to the nth resonance from Eq. (24) is somewhat larger and it is about 0.64 N ($n=1,2$). This value divided by the tube cross-section (0.78 cm^2 at $a=0.5$ cm) and multiplied by Q (say 5) gives us a pressure oscillation amplitude of 0.4 atm. It is probably too low to expect a considerable practical application of this resonator type. Compared with internal pressures $\rho_0c_0^2/(B/A+1)$ for the liquid metals, it is also too low for any nonlinear effects.

E. Solution: Open–open resonator

The general solution again has the form of Eq. (25) but $\cos k_1z$ is replaced by $\sin k_1z$, $\sin k_1z$ is replaced by $-\cos k_1z$; $k_1=\pi/L,\ 2\pi/L$ for the fundamental and second resonance, respectively. The effect of flexible tube walls and that of acoustic boundary layers are, as before, given by Eqs. (26) and (27), respectively. However, the radiation boundary condition (23a) for the flanged tube yields

$$\Omega=c_0k_1\left(1-\frac{8}{3\pi}\frac{a}{L}\right),\quad Q=\frac{\pi}{2}\frac{1}{(k_1a)^2} \tag{30}$$

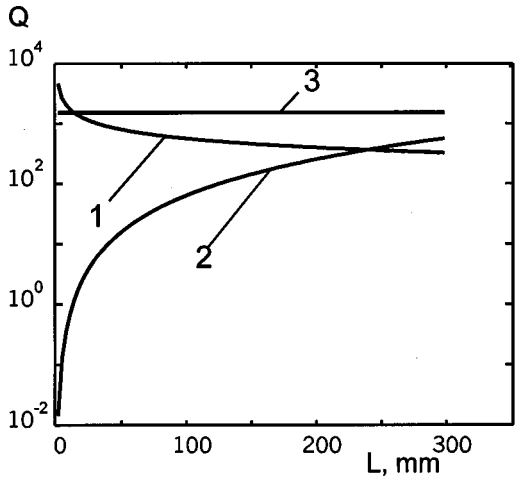


FIG. 5. Quality factor of the open–open resonator (liquid gallium) as a function of height L of the liquid metal column at $a=5$ mm. 1—effect of the boundary layer losses; 2—effect of open-end acoustic losses (linear theory); 3—effect of open-end hydrodynamic losses (nonlinear theory, $F_n=0.64$ N).

instead of Eq. (28). The resonant frequency agrees to main order with the calculation for the piston-driven acoustic resonator (Ref. 21, p. 202) if the driver presents a vanishing mechanical impedance. Figure 5 shows the quality factors given by Eqs. (27) and (30), respectively, as a function of height L of the liquid metal column. The resulting quality factor appears, in general, much greater than the quality factor for the open–closed resonator, but as L is considerably greater than the tube cross-section.

We check here the example given at the end of previous subsection and assume the same driving force of 0.64 N. For the open–open resonator, the pressure oscillation amplitude reaches 12.7 atm already for a tube with $a=5$ mm and $L=200$ mm, if the fundamental resonance is considered.

IV. NONLINEAR ANALYSIS

It is well known from both theory and experiment for open tubes that Eqs. (27) and (30) overestimate the quality factor at high oscillation amplitudes, whereas, at small amplitudes, they are close to reality.[40] The reason for such disagreement is the nonlinearity of the hydrodynamic flow, which drastically reduces the oscillation amplitude. For open resonators, there are two sources of nonlinearity: the nonlinearity of the flow within the tube and the boundary condition nonlinearity. The latter appears due to boundary layer separation from the tube wall at outflow (inflow), which is basically more important[40–43] than the familiar nonlinear distortion in the inner volume.

We apply the following nonlinear boundary condition[40] replacing Eq. (23a):

$$p(0,t)=-\rho_0\alpha|w(0,t)|w(0,t) \quad \text{at } z=0, \tag{31}$$

where α is a numerical factor on the order of unity. This type of nonlinear boundary condition is also known as Bernoulli's loss. Disselhorst and van Wijngaarden[40] discuss $\alpha=0$ at outflow and $\alpha=0.5$ or 1 at inflow for an unflanged tube with round or sharp edges, respectively. For the final result, it does not matter whether α is a step function or a constant: only its time-averaged value over a wave period is significant. Therefore, we set α to be a constant everywhere. Physically, this means an equal loss distribution between inflow and outflow.

Further treatment follows the standard template of nonlinear acoustics[44,45] (see also Refs. 30, 46). Dimensionless wave amplitude plays the role of a small but finite parameter. The flow in the tube is divided into forward p^+,w^+ and reverse p^-,w^- traveling waves, i.e.,

$$\begin{aligned} &p=p^++p^-, \quad w=w^++w^-,\\ &p^+=\rho_0c_0w^+, \quad p^-=-\rho_0c_0w^-. \end{aligned} \tag{32}$$

The real part of the force expansion Eq. (8b) (we assume for simplicity that $\mathrm{Im}\{f_n\}=0$) may be cast for the fundamental and the second resonance in the form

$$\begin{aligned} \mathrm{Re}\{f_ne^{j\omega t}\cos k_{1n}z\}&=\tfrac{1}{2}f_n\cos\omega(t-z/c_0)\\ &\quad+\tfrac{1}{2}f_n\cos\omega(t+z/c_0), \end{aligned} \tag{33}$$

where we have substituted $k_{1n}\cong\omega/c_0$ at resonance. Applying the multiple-scale asymptotic technique to the hydrodynamic equations of motion, it can be shown that both elementary traveling waves should satisfy the inhomogeneous simple wave equation (cf. Ref. 45)

$$\frac{\partial w^\pm}{\partial z}-\frac{n+1}{2c_0^2}w^\pm\frac{\partial w^\pm}{\partial\tau}=\pm\frac{f_n}{4\rho_0c_0}\cos\omega\tau \tag{34}$$

to second order. Here, $\tau=t\mp z/c_0$ is the retarded time. Equation (34) ignores the interaction between two oppositely directed waves. Fortunately, the nonlinear interaction between two counterpropagating waves is not cumulative over one period and contributes finally only a third-order correction.[47]

Substituting Eq. (32) into Eq. (31) and Eq. (23b), one obtains

$$\begin{aligned} w^+(0,\tau)-w^-(0,\tau)=-\frac{\alpha}{c_0}&|w^+(0,\tau)+w^-(0,\tau)|(w^+(0,\tau)\\ &+w^-(0,\tau)), \end{aligned} \tag{35a}$$

$$w^+(L,\tau)-w^-(L,\tau)=0. \tag{35b}$$

On the right-hand side of Eq. (35a) we can replace $w^-(0,\tau)$ by $w^+(0,\tau)$ to first order. This gives

$$w^+(0,\tau)-w^-(0,\tau)=-\frac{4\alpha}{c_0}|w^+(0,\tau)|w^+(0,\tau), \tag{36a}$$

$$w^+(L,\tau)-w^-(L,\tau)=0. \tag{36b}$$

Expansion into a Taylor's series and using Eqs. (34) yields

$$\begin{aligned} w^\pm(L,\tau)&=w^\pm(0,\tau)+L\frac{\partial w^\pm}{\partial z}(0,\tau)\\ &=w^\pm(0,\tau)+\frac{n+1}{2c_0^2}Lw^\pm(0,\tau)\frac{\partial w^\pm(0,\tau)}{\partial\tau}\\ &\quad\pm L\frac{f_n}{4\rho_0c_0}\cos\omega\tau. \end{aligned} \tag{37}$$

If we substitute this expression into Eq. (35b), the following relation is obtained ($\tau=t\mp z/c_0=t$ at $z=0$):

$$w^+(0,\tau)-w^-(0,\tau)=-\frac{n+1}{4c_0^2}L\frac{\partial}{\partial\tau}(w^{+2}(0,\tau)-w^{-2}(0,\tau))$$

$$-L\frac{f_n}{2\rho_0 c_0}\cos\omega\tau. \tag{38}$$

It is worth noting that the nonlinearity of the flow within the tube is represented by the first term on the right-hand side of Eq. (38). This term is cancelled to second order since we set $w^-(0,\tau)=w^+(0,\tau)$ to main order. Exactly the same is valid for the nonlinear correction to the two familiar expressions $p^+=\rho_0 c_0 w^+$, $p^-=-\rho_0 c_0 w^-$ from Eq. (32). Equation (38) is thus reduced to

$$w^+(0,\tau)-w^-(0,\tau)=-L\frac{f_n}{2\rho_0 c_0}\cos\omega\tau. \tag{39}$$

We see that the flow nonlinearity for open resonators only appears as a third-order effect[42,48] which is beyond the scope of our analysis. However, the boundary condition nonlinearity appears to second order already. Substitution of Eq. (39) into Eq. (36a) yields

$$w^+(0,\tau)=\frac{1}{2}\sqrt{\frac{Lf_n}{2\rho_0\alpha}}\sqrt{|\cos\omega\tau|}\,\mathrm{sgn}(\cos\omega\tau). \tag{40}$$

To within a constant factor, Eq. (40) coincides with the well-known van Wijngaarden's solution for the piston-driven acoustic resonator.[42,40]

Using Eq. (32), the maximum pressure amplitude P is found at antinodes

$$P\cong 2\rho_0 c_0 \max(w^+(0,\tau))=\sqrt{L\rho_0 c_0^2 f_n/2\alpha}. \tag{41}$$

The quality factor Eq. (24) is now calculated as follows:

$$Q=\frac{\pi}{4}\sqrt{\frac{2\rho_0 c_0^2}{\alpha L f_n}}=\frac{\pi}{2}\sqrt{\frac{\rho_0 c_0^2 a^2}{\alpha F_n}}. \tag{42}$$

The second expression involves the total resonance force F_n applied to the liquid metal volume. We have taken $\alpha=2$ in Eq. (42) to obtain the upper limit of the inflow–outflow losses and to account for experimentally observed additional losses (cf. Ref. 40).

Figure 5 plots the values of the quality factor Eq. (42) as a function of height L of the liquid metal column. Also shown are the quality factors predicted by the linear theory. In each case, the total driving force is 0.64 N (cf. example in previous section). For a 200-mm-long metal column ($a=5$ mm), the oscillation amplitude decreases from 12.7 to 11.5 atm due to the nonlinearity [the total quality factor Eq. (29) decreases by 10%]. This still represents a reasonably small correction. However, the nonlinear saturation effect increases rapidly with increasing oscillation amplitude.

V. DISCUSSION AND CONCLUSIONS

Figure 6 presents the contour plot of the maximum oscillation amplitude P (atm) for the open–open resonator (liquid gallium) as a function of height L of the liquid metal column and tube radius a. The total driving force takes the standard value of 0.64 N. Wall friction, acoustic radiation from the open end, and hydrodynamic inflow–outflow losses

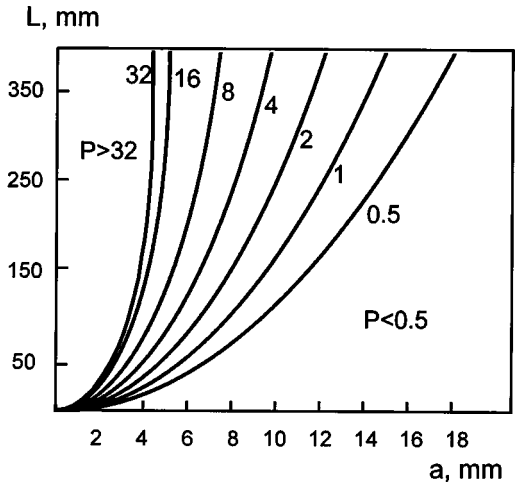


FIG. 6. Contour plot of maximum oscillation amplitude P (atm) for the open–open resonator (liquid gallium) as a function of height L of the liquid metal column and tube radius a. The total driving force is 0.64 N.

are simultaneously taken into account. We can see that this resonator is most effective at rather small tube radii of 2–4 mm. This conclusion is based on the nature of the acoustic and hydrodynamic open-end losses, which strongly depend on the tube radius.

One way to improve the resonator performance would be the use of a flattened rectangular tube under the assumption that the length of the cross-sectional rectangle is considerably greater than its width. Since the alternating current (cf. Sec. II) is allowed to flow along the larger rectangle side, the current path and the total force increase proportional to the side length. For example, if a rectangular tube with side ratio 1:4 is considered instead of a circular tube of the same cross-section, the total Lorentz force Eq. (7) becomes $\sqrt{\pi}\approx 1.8$ times larger. The radiation loss remains the same since it depends on the tube cross-section only. The boundary layer loss increases slightly depending on the cross-section perimeter.

Notice that the periodic current source is accompanied by a periodic Joule's heat source in the plane of two electrodes. The heat source itself can induce resonant oscillations much as periodic laser heating does in a resonance tube terminated by IR transparent windows.[49] In this case, we do not need the external magnetic field and the resulting Lorentz force. However, for good conductors, the heat source is estimated to be rather weak as a practical way of generating high-power resonance oscillations.

An important question beyond the scope of this paper is the effect of surface tension of the liquid metal on its free surface and the wetting effect of the tube wall on the quality factor of the resonator. We believe that, under presence of an acoustic field, there should be good wetting contact for liquid gallium and aluminum residing in quartz or ceramic tubes with weak surface roughness.

The following theoretical conclusions can be drawn based on the studies conducted in this paper:

(i) The resonant oscillation of a relatively small liquid metal column over a wide range of frequencies (5–40 kHz) can be excited by alternating Lorentz forces applied directly to the liquid metal.

(ii) The best performance is achieved for the open-open, half-wavelength resonator of circular or elongated cross-section. The presence of a rigid termination (open–closed quarter wavelength resonator) drastically reduces the quality factor of this resonator.

(iii) For the current of 35 A rms, pressure amplitudes of 10–20 atm can be expected at resonant oscillations in the frequency range of 5–40 kHz. Those conclusions are primarily related to liquid metal with rather high conductivity (liquid gallium or aluminum). Otherwise, Joule's heating plays an important role and additionally must be taken into consideration. Indeed, if the resonant tube is terminated by a local mechanical mass-spring system (for example, membrane fixed at the rim), then large pressure amplitudes can be obtained for the open–closed resonator as well. However, they are restricted to certain fixed eigenfrequencies.

[1] L. Bergmann, *Ultrasonics and Their Scientific and Technical Applications*, 3rd rev. ed. (United States Bureau of Ships, Washington, D.C., 1951).

[2] T. L. Tolt and D. L. Feke, ''Separation of dispersed phases from liquids in acoustically driven chambers,'' J. Chem. Eng. Sci. **48**, 527–540 (1993).

[3] R. E. Apfel, ''Sonic effervescence: A tutorial on acoustic cavitation,'' J. Acoust. Soc. Am. **101**, 1227–1237 (1997).

[4] G. I. Eskin, *Ultrasonic Treatment of Light Alloy Melts* (Gordon and Breach, Amsterdam, The Netherlands, 1998), Chaps. 2, 3.

[5] S. Asai, ''Recent activities on electromagnetic processing of materials,'' in *Proceedings of The Julian Szekely Memorial Symposium on Material Processing*, edited by H. Y. Sohn, J. W. Evans, and D. Apelian (Minerals, Metals & Materials Soc., Warrendale, PA, 1997), pp. 301–311.

[6] D. Apelian, ''Advances in metal treatment of aluminum and foundry alloys,'' in *Proceedings of the International Symposium on Light Metals 1997 Métaux Légers*, edited by C. M. Bickert and R. I. L. Guthrie (Canadian Institute of Mining, Metallurgy and Petroleum, Montreal, 1997), pp. 117–139.

[7] P. Pouly and E. Wuilloud, ''On the efficiency of in-line devices to clean the melt,'' in *Light Metals 1997*, edited by R. Huglen (The Minerals, Metals & Material Soc., Warrendale, PA, 1997), pp. 829–835.

[8] C. E. Eckert, ''Apparatus and method for ultrasonic detection of inclusions in molten metals,'' U.S. Patent 4,563,895 (1986).

[9] N. D. G. Mountford, I. D. Sommerville, A. Simeonescu, and C. Bai, ''Sound pulses enable the online visualization of liquid metal quality,'' in *Proceedings of the International Symposium on Light Metals 1997 Métaux Légers*, edited by C. M. Bickert and R. I. L. Guthrie (Canadian Institute of Mining, Metallurgy and Petroleum, Montreal, 1997), pp. 197–211.

[10] R. C. Stiffler, R. C. Wojnar, M. F. A. Warchol, L. W. Cisko, and J. M. Urbanic, ''Apparatus and method for ultrasonic particle detection in molten metal,'' U.S. Patent 5,708,209 (1998).

[11] R. T. Smith, G. M. B. Webber, F. R. Young, and R. W. B. Stephens, ''Sound propagation in liquid metals,'' in *The Properties of Liquid Metals*, edited by P. D. Adams, H. A. Davies, and S. G. Epstein (Taylor & Francis, London, 1967), pp. 515–522.

[12] G. M. B. Webber and R. W. B. Stephens, ''Transmission of sound in molten metals,'' in *Physical Acoustics*, Vol. IV-B, edited by W. P. Mason (Academic, New York, 1968), pp. 53–97.

[13] J. E. Hatch, Ed., *Aluminum: Properties and Physical Metallurgy* (American Society for Metals, Metals Park, OH, 1984), Chap. 1.

[14] T. Iida and R. I. L. Guthrie, *The Physical Properties of Liquid Metals* (Clarendon, Oxford, 1988).

[15] R. I. L. Guthrie, *Engineering in Process Metallurgy*, 2nd ed. (Clarendon, Oxford, 1992).

[16] G. T. Dyos and T. Farrell, Eds., *Electrical Resistivity Handbook* (Peter Peregrinus, London, 1992).

[17] I. S. Grigoriev, E. Z. Meilikhov, and A. A. Radzig, Eds., *Handbook of Physical Quantities* (CRC, Boca Raton, 1997).

[18] S. W. H. Yih and C. T. Wang, *Tungsten: Sources, Metallurgy, Properties, and Applications* (Plenum, New York, 1979).

[19] E. A. Almond, ''Deformation characteristics and mechanical properties of hard metals,'' in *Science of Hard Materials*, edited by R. K. Viswanadham, D. J. Rowcliffe, and J. Gurland (Plenum, New York, 1983), pp. 515–557.

[20] M. J. Lighthill, *Waves in Fluids* (Cambridge U.P., Cambridge, 1978).

[21] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics*, 3rd ed. (Wiley, New York, 1982).

[22] S. Makarov and E. Vatrushina, ''Effect of the acoustic boundary layer on a nonlinear quasiplane wave in a rigid-walled tube,'' J. Acoust. Soc. Am. **94**, 1076–1083 (1993).

[23] The term ''quasi-steady'' relies on the electromagnetic field. This means that we still calculate the current using an electrostatic solution (Laplace equation) multiplied by the periodically oscillating time factor. Otherwise, Maxwell equations should be solved to calculate the unsteady current distribution, which is not a very straightforward task. The background for such an assumption is the inequality $(a/\delta)^2 \ll 1$, where $\delta = \sqrt{2\rho_l/(\mu\omega)}$ is the skin depth layer in the liquid metal; $\mu \approx \mu_0 = 4\pi\,10^{-7}$ Wb/(A m) is the magnetic permeability. At $f = 10$ kHz, $\sigma \approx 2.6$ cm in liquid gallium or aluminum, which certainly applies for the quasi-steady solution, if tube radius does not exceed 1 cm.

[24] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vol. I, Vol. II (§ 10.3, p. 1263).

[25] W. R. Smythe, *Static and Dynamic Electricity*, 3rd ed. (McGraw-Hill, New York, 1967, 1968).

[26] Y.-H. Kim and S.-W. Kang, ''Green's solution of the acoustic wave equation for a circular expansion chamber with arbitrary locations of inlet, outlet port, and termination impedance,'' J. Acoust. Soc. Am. **94**, 473–490 (1993).

[27] E. G. Williams, ''On Green's functions for a cylindrical cavity,'' J. Acoust. Soc. Am. **102**, 3300–3307 (1997).

[28] Each term of the eigenmode expansion, Eq. (8), for the driving force contributes to the acoustic pressure in the tube. However, the magnitude of the contribution appears very different depending on the wave number of the mode k_{1n}. If $\omega/c_0 \approx k_{1n}$ then the nth eigenmode satisfies the resonance condition; all other modes are out of resonance. The repetition of the calculation below for a single nonresonant mode indicates that its contribution to the resulting pressure is at least Q times smaller (Q is the quality factor of the resonator) than the contribution of the resonant mode. Since we only consider the fundamental or second resonance, all nonresonant modes are of higher order. This additionally reduces the contributing pressures due to increased damping. As a result, we expect the total contribution of the nonresonant modes to be on the order of $1/Q$ as well. Precise calculations made for a test case (the wall friction is the only absorption mechanism) confirm this conclusion.

[29] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).

[30] S. Makarov and M. Ochmann, ''Nonlinear and thermoviscous phenomena in acoustics, Part II (a review),'' Acustica **83**, 197–223 (1997).

[31] P. M. Morse and K. U. Ingard, ''Linear acoustic theory,'' in *Encyclopedia of Physics. Volume XI/l Acoustics I*, edited by S. Fluegge (Springer-Verlag, Berlin, 1961), pp. 44–56, 119–121.

[32] M. C. Junger and D. Feit, *Sound, Structures, and Their Interaction* (MIT Edward Brothers, Cambridge, MA, 1972).

[33] H. Kraus, *Thin Elastic Shells* (Wiley, New York, 1967).

[34] T. C. Lin and G. W. Morgan, ''Wave propagation through fluid contained in a cylindrical, elastic shell,'' J. Acoust. Soc. Am. **28**, 1165–1176 (1956).

[35] M. El-Raheb, ''Acoustic propagation in finite length elastic cylinders. Part I: Axisymmetric excitation,'' J. Acoust. Soc. Am. **71**, 296–306 (1982).

[36] L. D. Lafleur and F. D. Shields, ''Low-frequency propagation modes in a liquid-filled elastic tube waveguide,'' J. Acoust. Soc. Am. **97**, 1435–1445 (1995).

[37] V. E. Nakoryakov, B. G. Pokusaev, and I. R. Shreiber, *Wave Propagation in Gas–Liquid Media* (CRC, Boca Raton, 1993).

[38] Regarding the first equality in Eq. (29), this is a trivial consequence of the perturbation theory. The second equality is more questionable. It was proved to be an asymptotically precise result if the different quality factors have considerably different magnitudes. In the following, exactly this situation will be studied. However, if the magnitudes are the same, Eq. (29) is only an approximation whose error might be on the order of 100%.

[39] Y. Shapira, ''Acoustic wave propagation in high magnetic fields,'' in *Physical Acoustics*, Vol. 5, edited by W. P. Mason (Academic, New York, 1968), pp. 1–58.

[40] H. M. Disselhorst and L. van Wijngaarden, ''Flow in the exit of open pipes during acoustic resonance,'' J. Fluid Mech. **99**, 293–319 (1980).

[41] U. Ingard and H. Ising, ''Acoustic nonlinearity of an orifice,'' J. Acoust. Soc. Am. **42**, 6–17 (1967).

[42] L. van Wijngaarden, ''On the oscillations near and at resonance in open pipes,'' J. Eng. Math. **2**, 225–240 (1968).
[43] R. G. Galiullin, A. Z. Murzakhanova, and I. P. Revva, ''Influence of absorption on the nonlinear oscillations of a gas in a pipe open at one end,'' Sov. Phys. Acoust. **36**, 545–547 (1990).
[44] O. V. Rudenko and S. I. Soluyan, *Theoretical Foundations of Nonlinear Acoustics* (Consultants Bureau, New York, 1977).
[45] V. E. Gusev, ''Build-up of forced oscillations in acoustic resonators,'' Sov. Phys. Acoust. **30**, 121–125 (1984).
[46] S. Makarov and M. Ochmann, ''Nonlinear and thermoviscous phenomena in acoustics, Part I (a review),'' Acustica **82**, 579–606 (1996).
[47] B. R. Seymour and M. P. Mortell, ''Nonlinear resonant oscillations in open tubes,'' J. Fluid Mech. **60**, 733–749 (1973).
[48] M. Ochmann and S. Makarov, ''Nonlinear and thermoviscous phenomena in acoustics, Part III (a review),'' Acustica **83**, 827–844 (1997).
[49] R. Raspet, B. Denardo, H. E. Bass, J. Brewster, and J. Kordomenos, ''Investigation of parametric drive of a longitudinal gas-filled resonance tube,'' J. Acoust. Soc. Am. **99**, 725–729 (1996).

Measurement of the stiffness coefficients of a viscoelastic composite material with laser-generated and detected ultrasound

S. Guilbaud and B. Audoin[a)]
Laboratoire de Mécanique Physique, CNRS UPRES A 5469, Université Bordeaux I, 351 cours de la Libération, 33405 Talence, France

(Received 4 February 1998; revised 9 December 1998; accepted 21 December 1998)

A point source-point receiver technique, based on laser generation and laser detection of acoustic waves, is used to measure the stiffness tensor of a composite material. Because of the dispersive nature of these materials, the acoustic signature provided in polymer matrix composites is difficult to interpret. A model for acoustic propagation in composites is developed to analyze the accuracy of the stiffness coefficients determined from the measured acoustic signals. The frequency dependence associated with viscoelastic rheology leads to the numerical implementation of a double inverse transform to calculate the surface displacement generated by a line impact. The model accurately predicts the focusing effects caused by anisotropy and the spreading of the signals caused by dispersion and attenuation. The predicted displacements are in good agreement with the result of experiments performed with a composite plate. The model is used to test a signal processing method based on wavelet transform to measure the group velocities of bulk waves propagating through the plate. Despite dispersion and echoes overlapping, the stiffness coefficients are identified from group velocities with good reliability. The process is then successfully applied to an actual viscoelastic composite material. © *1999 Acoustical Society of America.* [S0001-4966(99)00304-5]

PACS numbers: 43.35.Cg [HEB]

INTRODUCTION

The focusing of a laser beam on the surface of a medium has been recognized for years to be a powerful mean for generating acoustic waves.[1,2] When operating in the thermoelastic regime, i.e., at a low power level, it generates ultrasonic displacement through thermal expansion, and the source can be represented as a dipole strength, parallel to the surface of the specimen.[3] Alternatively, at high power level, the laser source operates in the ablation regime by vaporizing a small amount of surface material. It can then be modeled by a force normal to the surface of the specimen. The radiated field of such a source resembles a monopole radiating strongly in all directions from the source point.[4]

The laser-induced ultrasonic source is a broadband repeatable source which is readily amenable to operation in a scanning mode, and can also be used at high repetition rate with pulsed lasers. When it is coupled with an optical detection, it provides an appealing means to perform a nondestructive testing of materials and structures. This noncontact method is particularly well suited to analyze the mechanical behavior of new composite materials which are intended to be used at elevated temperature. However, it is only recently that the direct problem consisting in the calculation of the acoustic field radiated in any direction of an anisotropic elastic medium has been accurately analyzed. These works have been used to solve the inverse problem of the characterization of composite materials.[5] When dealing with polymer matrix composite materials an additional effort is required for both direct and inverse problems because of the viscoelastic behavior of the medium.

Section I is a brief overview of the main effects of anisotropy on acoustic propagation in viscoelastic materials. Their description enables us to interpret the intricate shape of the experimentally recorded signals. This section gives the preliminary explanatory matter for the following one, in which these phenomena are numerically represented.

The calculation of the displacement field in any direction of an elastic medium of unrestricted anisotropy has been solved only recently.[6–8] Viscoelastic anisotropic media have been the scope of works by Mal and Lih[9] who considered a source concentrated in space and narrow band in frequency. For an impulsive source, Weaver *et al.*[10] have implemented a numerical scheme to calculate the dynamic response of an elastic plate with the axis of transverse isotropy normal to the plate interface. It is adapted in Sec. II of the paper to the calculations of the normal displacements generated in any direction through a viscoelastic orthotropic plate, by a line source lying along a symmetry axis.

The first step when dealing with the measurement of the stiffness coefficients of an anisotropic material is to process the experimental waveforms to identify the group velocities of each generated mode. The arrival time measurement is a detection and localization problem that is treated in some works related to laser ultrasonics by the estimation of the time when peaks, identified as wave arrivals, show their maximum magnitude.[11,12] However, this method is doomed to failure for the localization of shear waves in most off epicenter waveforms. Other authors prefer to detect the maxima of the signal derivative which actually leads to locating the abrupt changes of the signal magnitude.[13] To these approaches, based on the extrema of either the interface displacement or its velocity, a more efficient method relying on the localization of the energy in the waveform[5] is preferred. More precisely, a time and frequency analysis based on a wavelet transform is considered in Sec. III. This transform is

[a)]Electronic mail: audoin@lmp.u-bordeaux.fr

well suited to find the location of singularities in the waveforms, since by decomposing signals into elementary building blocks that are well localized both in time and in frequency, it can characterize the local changes of the signals.[14]

In the second step, a numerical method is used to recover the stiffness coefficients from a set of group velocities. A double iterative inversion process is required, since there is no equation available that relates the group velocities to the elastic constants.[15] In Sec. IV, the process is applied starting with simulated signals, calculated with representative stiffness and viscosity coefficients. The difference between the identified stiffness coefficients and the reference values is analyzed before applying the identification process to the actual composite material.

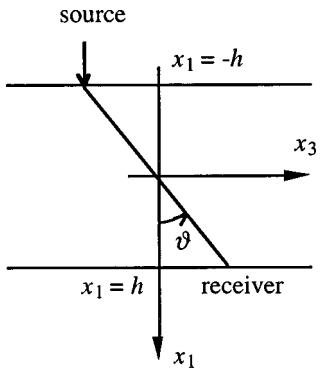


FIG. 1. Geometry of the problem in the plane (x_1, x_3).

I. ACOUSTIC WAVES GENERATED IN A VISCOELASTIC ANISOTROPIC PLATE

The device used for the noncontact generation and detection of the acoustic waves consists of two lasers: a Nd:Yag pulsed laser for generation (with or without frequency doubling depending on the amount of energy desired) and a continuous wave (CW) frequency doubled Nd:Yag laser for the detection.

The pulsed laser burst duration is about 10 ns. The beam is focused on the sample surface by using either a cylindrical or a spherical lens, providing a line or pointlike source, respectively. When an ablative line-source is desired, infra-red light emission at 1064 nm must be used with a medium burst energy output typically of 340 mJ. The collimated optical beam is focused providing a line whose length and width are about 6 cm and 0.1 mm, respectively. Otherwise, frequency doubled green light emission provides enough energy when an ablative point source is desired.

The CW laser is used as a Mach-Zehnder[16] interferometer to detect the normal displacement at the plate interface opposite to the source. Its bandwidth extends from 200 kHz to 18 MHz. A periscopic device permits us to move the impact location on the sample surface. The direction of observation is then defined by the angle ϑ between the source-receiver direction and the plate normal.

The anisotropic plate consists of a composite material made of carbon fibers embedded in a polyimide matrix which gives rise to the viscoelastic behavior of the specimen. The sample thickness is $2h = 6.3$ mm and the material density is $\rho = 1.7$ kg/dm^3. Because of its structure, the material has three perpendicular planes of symmetry. The direction of the fibers, axis 3, and the plate normal, axis 1, are principal directions. The remaining, axis 2, lies parallel to the sample surface (Fig. 1).

Let us now focus our interest on the acoustic waves observed when a line-source lying along the axis 2 is produced. The source to receiver direction is then included in the plane (1, 3). For a plane wave with a wave vector oriented in such a plane, only four coefficients of the fourth order stiffness tensor $\mathbf{C^e}$ interfere in the propagation equation.[17] The coefficients considered for illustration are given in Table I, using abbreviated subscripts notation. These values are chosen since they are close to those estimated for the composite material under the scope of this paper, by means of a more conventional ultrasonic assessment method that uses immersion devices.[18]

Three waves can propagate in this principal plane whose polarizations are quasi-longitudinal, quasi-transverse and pure-transverse. The latter has a polarization along axis 2 which is normal to the propagation plane. When the laser beam is focused along a line parallel to the direction 2, this pure traverse mode is not generated. The line-source generates transient divergent waves that propagate at group velocity. The phase and group velocities of acoustic waves in elastically anisotropic solids are, in general, not equal, even in the absence of dispersion and attenuation. The direction of group velocity, or energy flux, at any point of the slowness surface is parallel to the surface normal at that point. The phase slowness curves in plane (1, 3) of the composite material are shown in Fig. 2(a). The group velocities, or ray vectors, are calculated for points of the slowness curves associated with phase angles ϑ_p varying from 0° to 90° with a constant step. These group velocities are plotted in Fig. 2(b) as a polar function of the group angle ϑ, i.e., the direction of the normal to the slowness curves. Assuming the directivity pattern of the source to be cylindrical, the density of points along the curves in Fig. 2(b) is representative of the energy distribution.[19] Several comments can be drawn from these curves. Regarding the quasi-longitudinal mode (L), the shape of the group velocity curve is connected with the high value of the ratio C^e_{33}/C^e_{11}. Since the slowness curve is very flat around direction 3, the ray vectors appear very concentrated in this direction. Unfortunately, the group velocity of the waves propagating in this fiber direction is not experimentally accessible with this technique. Conversely, the energy of the quasi-longitudinal waves generated in the angular sector up to 70° from axis 1 is very weak. Concerning the quasi-transverse mode (T), the slowness curve is not entirely convex but it possesses regions where the curvature is negative. The group velocity curve then contains cuspical edges where the curve folds back on itself.[15] Within this folded region there are up to three group velocities in each direction. At the point of the slowness curve between two concave-convex regions, the curvature is zero. Therefore, the acoustic energy density is very high in the corresponding region of the ray curve. Nevertheless, seeing that the slowness curve does not show any axis of symmetry in the region where the curvature changes, the focusing observed at the

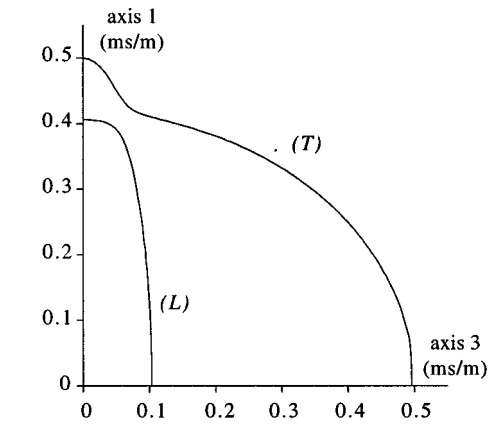
TABLE I. Stiffness coefficients (GPa) of the composite material.

	C_{11}^e	C_{33}^e	C_{13}^e	C_{55}^e
Used for simulations	9.7	150	1.3	6.4
Processing of calculated signals	10.0	151	3.9	6.4
Measurements	9.7±0.4	188±97	3.5±3.5	6.7±0.4

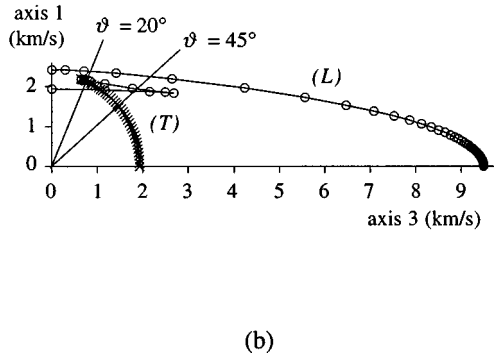
cuspidal edges is not symmetric. Then, the density of the energy concentrated in the cuspidal branch, denoted by crosses in Fig. 2(b), is high. Other factors, such as source directivity and polarization selectivity in the ablative generation process can also influence the directional dependence of the acoustic flux, but in a much less sensitive manner.[19]

These phenomena are now illustrated dealing with experimentally recorded waveforms. The normal displacement measured for an angle of observation ϑ of 20° is presented in Fig. 3(a). First, on account of the material attenuation, the waveform is smoothed in comparison with those measured after propagation through an elastic medium.[5] Moreover, since the group velocities of the quasi-longitudinal and quasi-transverse modes are very close to each other for this angle [see Fig. 2(b)], the corresponding waves are difficult to distinguish. Second, as discussed above, the acoustic energy of the quasi-transverse mode is much higher than that of the quasi-longitudinal mode for this angle. Consequently, despite the fact that the polarization of the quasi-transverse wave has a small component along axis 1, the amplitude of the associated contribution to the normal displacement is much higher than that of the quasi-longitudinal mode. This result is consistent with the acoustic energy focusing described above.

The experimental normal displacement for an angle ϑ of 45° is reported in Fig. 3(b). The same remark holds for the high amplitude of the contribution of the quasi-transverse mode to the normal displacement, provided by the energy distribution. The contribution (C) of only one branch of the

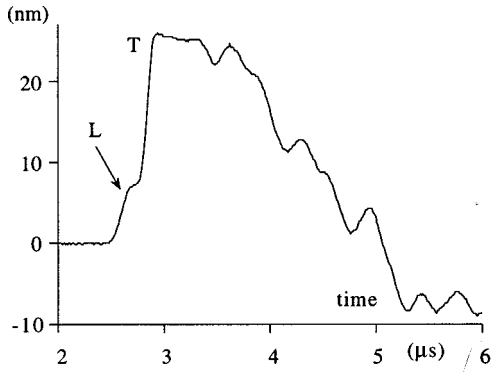


(a)

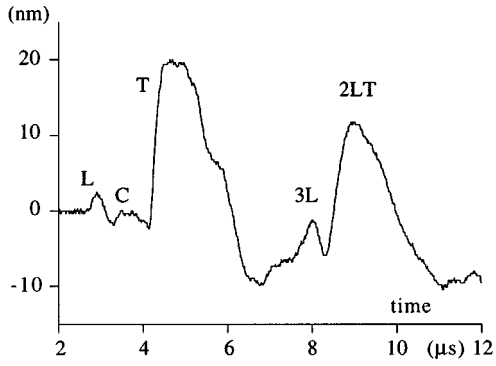


(b)

FIG. 2. (a) Phase slowness curves and (b) group velocity curves in the principal plane (1,3). The density of data (circles and crosses) illustrates the acoustic energy focusing. Data denoted by crosses point out the branch of the group velocity curve for which the focusing is the highest.



(a)



(b)

FIG. 3. Normal displacements for observation angles in the plane (1,3) of (a) ϑ=20° and (b) ϑ=45°.

cusp appears in the waveform. The displacement induced by the other branch is not detected since the associated energy is very low. Let us notice the later arrivals, denoted $(3L)$ and $(2LT)$, of waves propagating back and forth in the plate and reflected twice.

Finally, there is no difficulty in interpreting the measured waveforms, so long as the striking effects that anisotropy gives rise to are clearly kept in mind. However, because of the waveforms spreading and the signals overlapping, the accurate measurement of the group velocities appears to be a nontrivial matter. It is the purpose of the following section to present a calculation of the interface displacement that takes both the effects of anisotropy and of viscoelasticity into account.

II. CALCULATION OF THE WAVEFORMS

To a large extent, the success of laser ultrasonics has been the researcher's ability to correctly predict the time dependent evolution of the displacement waveform resulting from laser irradiation. Instead of the usual evaluation of the wave amplitude in the time domain by the Cagniard method,[8,15,20] it is hereafter evaluated in the frequency domain. The method implemented has first been used to calculate the response of an elastic material with restricted symmetry.[7,10] Similar modeling has been introduced to represent the waveforms generated in isotropic viscoelastic materials.[21] Due to the frequency dependence associated with the viscoelastic rheology, the method is especially suited to describe the propagation in such media. In this section, the calculation of the normal displacement generated by a line-source through a viscoelastic orthotropic plate is presented.

A. The surface line load problem

The medium is thought to be a homogeneous and infinite plate of thickness $2h$ and density ρ, with an orthorhombic symmetry. The axes of the Cartesian reference (x_1,x_2,x_3) are defined by the three planes of symmetry of the medium, with the x_1 axis normal to the plate. The normal displacement response at a point of the surface $x_1=h$ is sought.

In order to describe the viscoelastic behavior of the material, complex components of the stiffness tensor $\mathbf{C}$ are considered. It is classically assumed that the imaginary parts of $\mathbf{C}$ are linear functions of the angular frequency ω.[17] The fourth order tensor $\boldsymbol{\eta}$ is thus introduced such that

$$\mathbf{C}=\mathbf{C^e}+j\omega\boldsymbol{\eta}. \tag{1}$$

A line source along a principal direction, axis 2, is considered. Due to the geometry of the problem, there is no loss of generality in assuming all vectors to be expressed in the plane (1, 3). The displacement field $\mathbf{u}=(u_1,u_3)$ satisfies the following propagation equations[17]

$$\rho\frac{\partial^2 u_1}{\partial t^2}=C_{1jkl}\frac{\partial^2 u_k}{\partial x_l\partial x_j},$$
$$\rho\frac{\partial^2 u_3}{\partial t^2}=c_{3jkl}\frac{\partial^2 u_k}{\partial x_l\partial x_j}. \tag{2}$$

In Eqs. (2) and in the following ones, the repeated indices take values 1 and 3 only.

The two-dimensional Fourier transform $\mathbf{U}=(U_1,U_3)$ of the displacement field over coordinate x_3 and time t is now considered. The components of the displacement at a given position and time are

$$u_i(x_1,x_3,t)=\frac{1}{(2\pi)^2}\int\int_{-\infty}^{+\infty}U_i(x_1,k_3,\omega)\times\exp\{-j(k_3x_3-\omega t)\}dk_3d\omega, \tag{3}$$

where k_3 denotes the component of the wave number in direction 3. The variables ω and k_3 are Fourier dual forms of t and x_3, respectively. The set of equations (2) is now written in terms of Fourier transforms of $\mathbf{u}$ with respect to x_3 and t. Equations (2) then convert to

$$(\rho\omega^2-k_3^2C_{55})U_1+C_{11}\frac{\partial^2U_1}{\partial x_1^2}-jk_3(C_{55}+C_{13})\frac{\partial U_3}{\partial x_1}=0, \tag{4}$$
$$(\rho\omega^2-k_3^2C_{33})U_3+C_{55}\frac{\partial^2U_3}{\partial x_1^2}-jk_3(C_{31}+C_{55})\frac{\partial U_1}{\partial x_1}=0.$$

The complex component in direction 1 of the wave vector is denoted by k. Its imaginary part represents the exponential decay of the wave amplitude in direction 1. By expressing $\mathbf{U}$ under the new form $\mathbf{U}=\overline{\mathbf{U}}e^{-jkx_1}$, the system (4) reduces to a linear and homogeneous system written as follows in matrix form:

$$\begin{pmatrix}\rho\omega^2-k_3^2C_{55}-k^2C_{11} & -kk_3(C_{13}+C_{55})\\ -kk_3(C_{13}+C_{55}) & \rho\omega^2-k_3^2C_{33}-k^2C_{55}\end{pmatrix}\begin{Bmatrix}\overline{U_1}\\ \overline{U_3}\end{Bmatrix}=\begin{Bmatrix}0\\0\end{Bmatrix}. \tag{5}$$

Solving the characteristic equation of (5) leads to two solutions for k^2. The solutions $k=\pm k_l$ and $k=\pm k_t$, refer to the quasi-longitudinal and quasi-transverse propagation modes, respectively. The sign denotes the forward or backward traveling waves. The polarization vector $\mathbf{U}$ of the relevant k value is also determined from Eq. (5).

To calculate the amplitude of each mode contribution, the boundary conditions are considered. At the front interface $(x_1=-h)$, the impulsive laser source is described by a force $\mathbf{F}=(F_1,0,0)$ whose time shape is a delta function. Besides, free-loaded conditions are assumed on the rear interface $(x_1=h)$. The boundary conditions are then written as follows:

$$C_{1jkl}u_{k,l}\big|_{x_l=-h}=F_j\delta(x_3)\delta(t),$$
$$C_{1jkl}u_{k,l}\big|_{x_l=h}=0. \tag{6}$$

Just as previously, the double Fourier transform of Eqs. (6) is considered. This leads to a new system of partial differential equations. The transformed boundary conditions are

$$C_{11}\left.\frac{\partial U_1}{\partial x_1}\right|_{x_1=-h} - jk_3C_{13}U_3(x_1=-h)=F_1,$$
$$C_{55}\left(-jk_3U_1(x_1=-h)+\left.\frac{\partial U_3}{\partial x_1}\right|_{x_1=-h}\right)=0, \tag{7}$$

at the front interface, and

$$C_{11}\left.\frac{\partial U_1}{\partial x_1}\right|_{x_1=h} - jk_3C_{13}U_3(x_1=h)=0,$$
$$C_{55}\left(-jk_3U_1(x_1=h)+\left.\frac{\partial U_3}{\partial x_1}\right|_{x_1=h}\right)=0, \tag{8}$$

at the initially quiescent rear interface.

The symmetric and antisymmetric solutions with respect to the variable x_1,

$$\mathbf{U}^{\mathrm{sym},k}=\begin{Bmatrix}\overline{U_1^k}\cos(kx_1)\\ j\overline{U_3^k}\sin(kx_1)\end{Bmatrix} \text{ and } \mathbf{U}^{\mathrm{ant},k}=\begin{Bmatrix}\overline{U_1^k}\sin(kx_1)\\ -j\overline{U_3^k}\cos(kx_1)\end{Bmatrix}, \tag{9}$$

and formed in order to express the general form of the solution as a linear combination of these modes. Their respective amplitudes are denoted by the coefficients $A^{\mathrm{sym},k}$ and $A^{\mathrm{ant},k}$ such that

$$\mathbf{U}=\sum_{k=k_l,k_t}(A^{\mathrm{sym},k}\mathbf{U}^{\mathrm{sym},k}+A^{\mathrm{ant},k}\mathbf{U}^{\mathrm{ant},k}). \tag{10}$$

The system of four equations (7) and (8) is now written using the solution form given by Eq. (10). A combination of pressure and shear stress boundary conditions provides two uncoupled systems for $A^{\mathrm{sym},k}$ and $A^{\mathrm{ant},k}$, respectively. The following equations are obtained:

$$\sum_{k=k_l,k_t}A^{\mathrm{sym},k}(kC_{11}\overline{U_1^k}-k_3C_{13}\overline{U_3^k})\sin(kh)=\frac{F_1}{2},$$
$$\sum_{k=k_l,k_t}A^{\mathrm{sym},k}(k\overline{U_3^k}-k_3\overline{U_1^k})\cos(kh)=0, \tag{11}$$

for the symmetric modes, and

$$\sum_{k=k_l,k_t}A^{\mathrm{ant},k}(kC_{11}\overline{U_1^k}-k_3C_{13}\overline{U_3^k})\cos(kh)=\frac{F_1}{2},$$
$$\sum_{k=k_l,k_t}A^{\mathrm{ant},k}(k_3\overline{U_1^k}-k\overline{U_3^k})\sin(kh)=0, \tag{12}$$

for antisymmetric modes.

Let us now focus our interest on the calculation of the integral in Eq. (3). When dealing with an elastic medium, the integrand shows discontinuities for particular k_3 values. They correspond to poles associated with the zeroes of the dispersion equation that describe the guided waves in the plate. Thus, the evaluation of the integral appears to be not consistent with the Fourier transformation. For an exact calculation of the displacement field, it should be carried out in the complex plane of the variable k_3. In this case, the evaluation is performed by a contour integration, and with the aid of the Cauchy residue theorem. See Refs. 15, 22, and 23 for discussion of the general approach and Ref. 8 for applications to the case of an anisotropic plate. Since this method includes a change of variable[22] which involves time, it is not consistent with the frequency dependence involved by the propagation in a viscoelastic medium. A numerical integration method should therefore be applied.

TABLE II. Viscosity coefficients (kPa.s) considered for the simulations.

η_{11}	η_{33}	η_{13}	η_{55}
0.043	0.4	0.016	0.015

For each value of the angular frequency ω, the integral on the real axis of the variable k_3 is calculated by means of the method used by Weaver *et al.*[10] In this scheme, the Fourier transform is generalized by replacing ω by a complex variable $\omega - j\delta$ with a small, constant and imaginary part δ. With this change of variable, Eq. (3) becomes

$$u_i(x_1=h,x_3,t)=\frac{1}{(2\pi)^2}\exp(\delta t)\int_{-\infty}^{+\infty}\left\{\int_{-\infty}^{+\infty}U_i(x_1=h,k_3,\omega-j\delta)\times\exp(-jk_3x_3)dk_3\right\}\exp(j\omega t)d\omega, \tag{13}$$

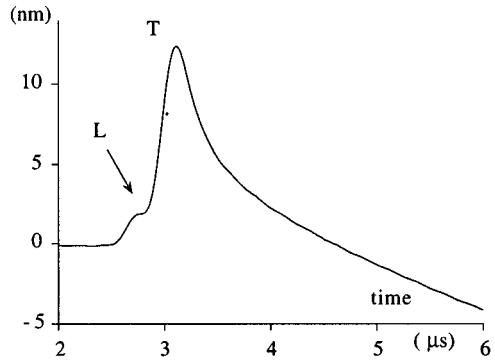
where the value of x_3 corresponds to the relative position of the line-source and of the point-receiver. The advantage of the method is that the integrand is a nonsingular function that may now be integrated numerically.

Let us notice that in the case of propagation in a viscoelastic medium, the poles of the integrand are moved off the real k_3 axis, and the asymptotic discontinuities are smoothed.[9] The integration may therefore be performed directly, without the need of introducing a small auxiliary parameter δ. However, in the case of a low viscosity or when the poles occur for low k_3 values, the integrand behavior remains nearly singular. A very short integration step or an adaptive integration method is required which is not cost effective. Consequently, when dealing with such a medium, the use of the δ method remains especially well suited in terms of numerical efficiency and accuracy.

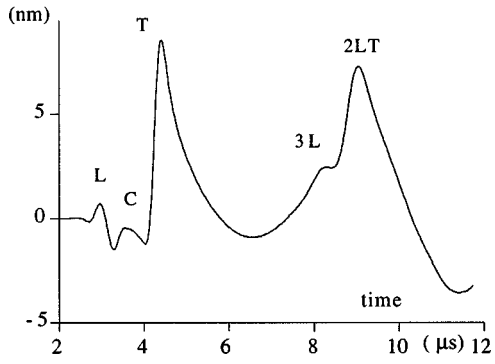
As a general remark, let us emphasize that, although the inverse transform calculation is restricted in this paper to the normal component u_1 of the displacement at the rear interface, it can be extended to the calculation of each component of the displacement at either the rear or front interface. This could be attractive when the detection of in-plane displacements is performed[24] and when surface waves are sought. Moreover, the method may be used to calculate the response of the plate to a thermoelastic or to a pointlike source. The latter source implies additional computational burden that is not necessary to illustrate the results presented below.

B. Comparison with experimental results

The model is now used to calculate the normal displacements corresponding to the experimental signals shown in Fig. 3. The viscosity coefficients considered to model the composite material are given in Table II. These values are



(a)



(b)

FIG. 4. Calculated normal displacements for observation angles in the plane (1,3) of (a) $\vartheta=20°$ and (b) $\vartheta=45°$. The magnitude of the source is $F_1=1$ N.μs.m^{-1}.

chosen since they are close to those measured by ultrasonic means for a similar composite material.[25]

The signals calculated for the observation angles ϑ of 20° and 45° are presented in Fig. 4, for a load magnitude $F_1=1$ N.μs.m^{-1}. The sampling frequency chosen for these simulations is similar to that of the experimental waveforms. Before turning our attention to a comparison with experimental normal displacements, we note that the interferometric probe processing includes a bandpass filter with a low cutoff frequency of 200 kHz. This technological limit leads the detected signal to be convoluted with a low frequency modulation. A numerical correction is then performed to represent this filtering effect.

The spreading of the calculated waveforms due to dispersion is consistent with the amplitudes and shapes observed for experimental signals shown in Fig. 3. The arrival times of the echoes associated with each mode are also identical. Concerning the calculated waveform for an angle ϑ of 20°, the overlapping of the longitudinal (L) and transverse (T) contributions and the relative amplitude of these two echoes are in very good agreement with the experimental signal. For $\vartheta=45°$, one observes in Fig. 4(b) the blurred shape of the cusp contribution (C). The calculations, taking into account both the effects of viscoelasticity and anisotropy, allow us to accurately represent this phenomenon. The later arrivals ($3L$) and ($2LT$), which correspond to waves propagating back and forth in the plate and reflected twice, are also noticeable. For both angles, discrepancy is observed between the calculated and experimental waveforms after the quasi-transverse wave (T) arrival. It is likely due to the line distribution of the laser energy which is not perfectly uniform on the line-source as assumed in the model. Other phenomena, such as diffraction by the fibers and the finite size of the source, may influence the waveforms, but in a much less sensitive manner.

Nevertheless, the effects of anisotropy and of viscosity that provide the multiple arrivals and the spreading of the signals are well predicted by the calculation method. In the following section, the waveforms generated by a line-source in various directions through the anisotropic viscoelastic medium are calculated. With such a tool in hand, an efficient signal processing is presented, which permits us to recover the group velocities associated with the generated waves.

III. PROCESSING OF THE SIGNALS

Because of dispersion and the subsequent spreading of the waveforms, the arrival time of the generated acoustic modes is difficult to estimate. Moreover, in view of the signals overlapping, see Figs. 3(a) and 4(a), there is a requirement for a high resolution processing. An efficient measurement of the group velocities is therefore under the scope of this section.

The approach is based upon a time-frequency analysis performed by means of wavelet transform.[26,27] Whereas classical Fourier analysis provides global information on the signal in the frequency domain, the wavelet transform enables us to investigate the signal changes versus frequency and time parameters. The signal $s(t)$ is no longer seen as a sum of unbounded $e^{j\omega t}$ functions but as a more complete two-dimensional decomposition, with respect to both time and frequency bounded functions.

To this aim, the recorded signal $s(t)$ is extended in the complex plane with the signal in quadrature, obtained with the Hilbert transform of $s(t)$ itself. The so-obtained complex signal $z(t)$ is the analytic signal of $s(t)$.[28]

Let us now consider an analytic signal $\varphi(t)$, bounded in the time and frequency domains, with a central angular frequency ω_0. This ''mother wavelet'' φ is used to generate a wavelet family whose components are denoted $\varphi_{a,b}$, where a and b are real and positive parameters. The scale parameter a is a dilatation factor that tunes the central frequency of the wavelet $\varphi_{a,b}$. The parameter b is a time translation factor that centers the wavelet on the desired time window. Let us note that the $\varphi_{a,b}$ functions are also time bounded and frequency bounded. Any mother wavelet generates an admissible wavelet family if: first, the average value of the function $\varphi(t)$ is zero, and second, an energy conservation law[29] is satisfied in order to ensure that there is no loss of information in the inverse transformation. Several kinds of families have been built, associated to different mother wavelets that fit the con-

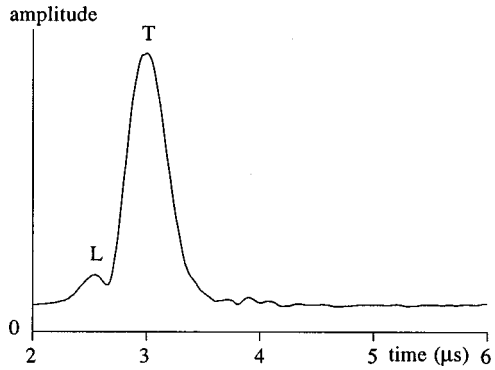


FIG. 5. Modulus of the wavelet transform of the signal in Fig. 4(a) (arbitrary units on a linear scale). The central frequency of the wavelet used is 5 MHz.

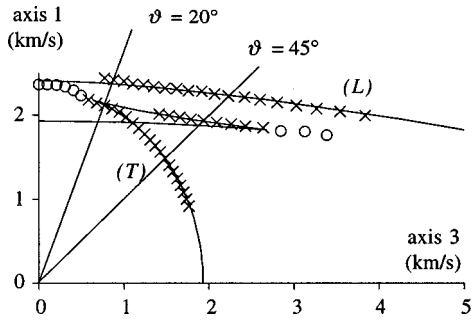


FIG. 6. The solid lines are group velocities calculated from reference values of stiffness coefficients. Data (circles and crosses) are recovered from simulated signals. Data denoted by circles illustrate the effect of diffraction by the cusp edges.

sidered problem. In this paper, the second derivative of a Gaussian function is chosen.

The continuous wavelet transform is defined by the scalar product of the analytic signal z with the wavelet $\varphi_{a,b}$. For a fixed dilatation parameter a, the transformation is the convolution of the analytic signal by the frequency filter centered at $\omega_0/2\pi a$. Consequently, the modulus of the complex coefficients $C_{a,b}$ provides the distribution of the signal energy in the time-frequency plane.

There is a certain emphasis, in the concerned literature, on the use of the wavelet transform to detect singularities in signals. Recently, Mallat and Hwang[14] have demonstrated that the Lipschitz exponent, which gives an indication of the differentiability of the signal, can be estimated by means of the evolution of the wavelet transform across scale. Moreover, it is proved in their paper that all singularities of $s(t)$ can be located by following the modulus maxima lines when the scale goes to zero. This convergence is of particular interest to us in seeing that the desired information appears located at low values of a. The position of the singularities is thus obtained with high central frequency wavelets for which the resolution requirements may be satisfied. For the purpose of illustration, let us perform the convolution of the signal in Fig. 4(a), with a wavelet whose central frequency is 5 MHz. This frequency is the highest beyond which the signal spectrum vanishes, i.e., information about the signal remains at this frequency and below. The modulus of the wavelet transform is plotted in Fig. 5. Despite the overlapping of the two echoes in the time domain, the quasi-longitudinal and quasi-transverse modes appear clearly separated in Fig. 5 inasmuch as two energy density maxima can be detected.

Several signals are now calculated by means of the numerical scheme detailed in Sec. II for observation angles varying from 0° to 62.5° with a 2.5° step. Each signal is processed by means of a wavelet whose central frequency is 5 MHz, and the locations of the energy density maxima are extracted. The associated velocities are denoted by either crosses or circles in Fig. 6. The solid lines represent the group velocities calculated with the real part of the stiffness coefficients used for the simulation of the signals. Their values are reported in Table I. The data denoted by circles are not related to any ray vector. They result from the proximity of the cuspidal edges and from the high energy distribution in this region where the ray curve folds back on itself. This acoustic energy focusing was pointed out in Sec. I. This phenomena, providing wavefronts that cannot be predicted by ray construction from phase slowness curve, is called diffraction by the cusp edges. It has yet been observed by several authors who are well referred in Ref. 30. Since this diffraction is accurately represented by the calculations, the arrival times of these wavefronts are obtained from the process.

The agreement between the recovered data denoted by crosses and the solid lines representing reference values shows the reliability of the method to recover group velocities. As a result, in addition to the possibility to resolve overlapped echoes, the location of the maxima of the modulus of the wavelet transform provides an accurate measurement of the group velocities. Moreover, due to the diffraction by the cusp edges, data that cannot be associated with any ray arrival are measured. In the remaining section, it is shown that the process is consistent with the inverse problem of the measurement of stiffness coefficients from detected waveforms.

IV. MEASUREMENT OF THE STIFFNESS TENSOR OF THE COMPOSITE MATERIAL

Four stiffness coefficients only can be measured by scanning the line source along a principal direction normal to it. Their identification is performed in a two-stage process. First, signal processing provides the measurement of group velocities for a set of observation angles. Second, a numerical algorithm is used to identify the stiffness coefficients from all the measured velocities. The numerical process for stiffness coefficients recovering from group velocities is based upon two analytical forms. A detailed description can be found in Ref. 31.

The first form follows from the propagation equations,[17] written with suitable coordinates

$$\mathcal{F}(t_0, p, \mathbf{X}, \vartheta) = 0, \tag{14}$$

where t_0 is the arrival time of a wave propagating at group velocity, ϑ denotes the direction of observation, and p stands for the projection on the interface of the phase slowness in direction ϑ_p, see Sec. I. In Eq. (14), $\mathbf{X}$ stands for the unknown coefficients vector. When a line source along direction 2 is used the components of $\mathbf{X}$ are $(C^e_{11}, C^e_{13}, C^e_{33}, C^e_{55})$.

The second analytical form expresses the geometric relationship between group velocities and slowness curves. It is written with the same variables as Eq. (14):

$$\mathcal{G}(t_0, p, \mathbf{X}, \vartheta) = 0. \tag{15}$$

The system, Eqs. (14) and (15), is then reversed in order to identify the optimal coefficients. The principal lies in the minimization of the L^2 norm of the functions $(\mathcal{F}(t_{0n}^{\exp}, p_n, \mathbf{X}, \vartheta_n^{\exp}))_N$, for N measured values of group times $t_{0n}^{\exp}$, and corresponding to N observation directions $\vartheta_n^{\exp}$. For each step of the variable $\mathbf{X}$, slownesses p_n are taken as parameters and calculated from the minimization of the norm of the functions $(\mathcal{G}(t_{0n}^{\exp}, p_n, \mathbf{X}, \vartheta_n^{\exp}))_N$. The algorithm is very robust in terms of convergence, and the area of attraction is very large, meaning that the initialization is almost with no importance. Moreover, this identification algorithm can take into account several velocities for one mode in a given direction. Other approaches[12] have been developed for this identification problem for which such multiple contributions could not be introduced in the inversion algorithm. These data are important for the identification process, since they are highly sensitive to the anisotropy of the medium.

By analyzing the residual squares of Eq. (14), a 90% confidence interval on the identified stiffness coefficients can be calculated.[32] It is based upon the analysis of the scattering of the experimental results around the exact solution. These confidence intervals quantify the identification quality from a set of group velocities. Still, it does not allow us to calculate an exact error because systematic errors on velocity measurements are not taken into account.

A. Simulations

The set of signals, calculated by means of the numerical scheme introduced in Sec. II, has been analyzed using the wavelet transform method described in Sec. III. The group velocity data, shown in Fig. 6, are now introduced in the numerical process for the recovery of the stiffness coefficients. On account of comments made in Sec. III, concerning the diffraction by the cusp edges, the data denoted by circles are not considered for the identification.

The recovered values are reported in Table I. They can be compared with reference values, used for the calculation of the waveforms. Several comments can be drawn from this comparison.

First, the recovered values of the coefficients C_{33}^e and C_{55}^e are very close to their reference counterparts. Thus, the measurement of these stiffness coefficients shows a good reliability. Second, discrepancies are observed for the coefficients C_{11}^e and C_{13}^e. These discrepancies result from systematic errors that arise from the effect of either dispersion or echoes overlapping. At this point, the question naturally arises as to what is the error introduced by each phenomenon. The ultrasonic dispersion observed when laser techniques are used with viscoelastic materials has been extensively analyzed for the case of an isotropic material.[21,33] The purpose is to evaluate the sensitivity of the stiffness coefficients' measurement to dispersion, for signals such that echoes do not overlap. To this aim, the waveform calculated for an observation angle of 0° is processed. However, the diffraction by the cusp edges observed in the plane (1, 3) and illustrated in Fig. 6 may influence the time measurement for angles down to 0°. The waveform is therefore calculated with a line source lying in direction 3 by introducing stiffness coefficients that account for the isotropy of the plane (1, 2). The resulting signal is processed by means of a 5-MHz central frequency wavelet and the arrival time of the quasi-longitudinal mode is calculated. An estimated value of 9.72 GPa is deduced for the stiffness coefficient C_{11}^e. The discrepancy between the recovered data and the reference value is smaller than the discrepancy observed for the values of the coefficient C_{11}^e given in Table I. This result allows us to neglect the effects of dispersion for a standard composite viscosity. In other words we consider that the changes of velocity with frequency are not sensitive, in regards to the stiffness identification accuracy. Consequently, the systematic error that provides the shift of the coefficients C_{11}^e and C_{13}^e results chiefly from echoes overlapping. The sensitivity of the signal processing to this overlapping may be minimized by using orthogonal wavelets.[27] Such developments are in progress and will not be presented in this paper.

B. Experiments

A set of signals is recorded from transmission experiments through the composite plate. The scanning of the front interface provides observation angles ranging from 0° to 65° with a 2.5° step. For the purpose of experimental convenience of measurements made at elevated temperatures,[34] a spherical lens is used. Thus, the source has a pointlike shape. Consequently the scattering by the source is no longer a two-dimensional problem whereas the scanning of the surface still leads to the investigation of the principal plane (1, 3). The quasi-longitudinal (L) mode, the quasi-shear (T) mode, and the associated cusp (C) previously observed are still generated. They propagate in any direction included in that plane of symmetry, with the same group velocity as the waves generated with a line source. Due to the three-dimensional scattering, some additional acoustic rays may appear. They are named the out of plane contributions, since they are associated with wave vectors that do not belong to the plane of observation.[35] Such waves have been detected in silicon[5,13] and zinc crystals.[6] The conditions on the stiffness coefficients for the observation of such acoustic waves have been established by Maris.[36] It can be shown that unlike these crystals, the quasi-transverse mode (T) of the composite material does not carry any out of plane contribution towards the plane (1, 3).

The contributions of the second shear mode, whose polarization is purely transverse in the plane (1, 2), must also be discussed in detail. From the generation point of view, the problem is no longer invariant with respect to direction 3, thus no reason exists for that mode not to be generated. However, due to the sensitivity of the detection device to normal displacement, a transverse wave propagating in the plane (1, 2) with its polarization along the direction 3 cannot be observed. Moreover, comments about the shape of the relevant slowness surface would also illustrate that no out of plane contributions occur in the scanned plane.

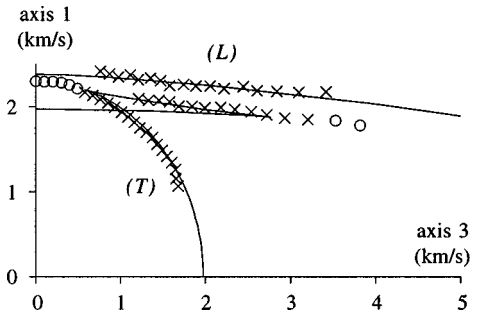


FIG. 7. Group velocities recovered from experimental signals measured for observation directions in the plane (1,3). Data denoted by circles illustrate the effect of diffraction by the cusp edges. The solid lines show the optimal fit of data denoted by crosses.

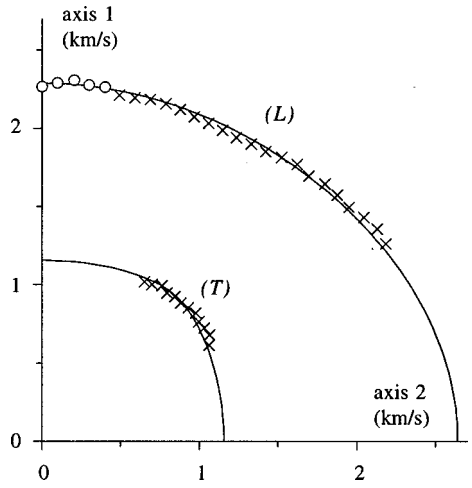


FIG. 8. Group velocities recovered from experimental signals measured for observation directions in the plane (1,2). The solid lines show the optimal fit of data denoted by crosses. Owing to the diffraction by the cusp edges shown in Fig. 7, data labeled by circles are not considered for stiffness identification.

Finally, due to the symmetry of the material, the detected waves are only the waves that are observed when a line source oriented along the direction 2 is used.

The set of experimental signals has been processed using the wavelet transform method described in Sec. III. The group velocity data are shown in Fig. 7. Because of the experimental noise, the signal to noise ratio in the wavelet bandwidth around 5 MHz is smaller when experimental signals are processed. Consequently, the data shown in Fig. 7 are more scattered than the velocity data in Fig. 6, which are processed from simulated signals. For the same reason, the group velocity of the quasi-longitudinal mode could not be measured at large angles for which the wave amplitude is weak.

The diffraction by the cusp edges provides undesirable data, and a bias occurs in the velocity of the quasi-longitudinal mode near epicenter. The effect is similar to that observed in Fig. 6. On account of comments made in Sec. III, the data marked by circles are not considered for the identification.

The recovered stiffness coefficients, as well as their associated 90% confidence intervals, are listed in Table I. The low confidence interval associated with the identification of C_{11}^e and C_{55}^e shows the reliability of their recovering from the group velocity data.[32] The coefficient C_{33}^e mostly influences the velocity in direction 3 which is not of experimental access. Because of the shape of the velocity curves, a small scattering in the velocity data provides the high 90% confidence interval associated with this coefficient identification. The confidence interval relative to C_{13}^e is also quite large because, as analyzed in Sec. IV A, this parameter identification is highly sensitive to the scattering of the velocity data. One must note that the choice of the data selected near the cusp edges, for the purpose of identification, influences the result, especially for C_{33}^e and C_{13}^e measurements. These sensitivities are illustrated by the large confidence intervals associated with their identifications. Moreover, the processing of the calculated waveforms has underlined the fact that echoes overlapping may induce errors in the identification of the coefficients C_{11}^e and C_{13}^e.

In order to achieve the stiffness coefficients identification, attention is turned to group velocities measured by scanning the point source along direction 2. The results are shown in Fig. 8. Owing to the quasi-isotropy of this principal plane of symmetry, multiples arrivals of the quasi-transverse mode are not observed. Nevertheless, since the source has a pointlike shape, the mixing of the quasi-longitudinal mode with the wave diffracted by the cusp edges actually occurs near epicenter in this plane (1, 2), as well as in the plane (1, 3). The corresponding data denoted by circles in Fig. 8 are not considered in the identification process.

Results of the stiffness coefficients identification are reported in Table III. Because of the quasi-isotropy of the principal plane (1, 2), the focusing-defocusing effects described in Sec. I and observed during scanning in the principal plane (1, 3) do not occur. The acoustic energy distribution is quasi-isotropic in the plane (1, 2). Consequently, the angular sector of experimental investigation is greater in this plane (1, 2) than in the plane (1, 3) under consideration above. The identification of the stiffness coefficient C_{22}^e then exhibits a better reliability than that of the coefficient C_{33}^e, as quantified by their relative confidence interval. Small discrepancies are observed between the two measurements of the coefficient C_{11}^e, see Tables I and III. This disagreement may arise from the systematic shift pointed out in Sec. III, and from variations of material properties from one sample to another.

V. CONCLUSION

The stiffness coefficients of a viscoelastic composite material have been measured by means of a noncontact experimental device that includes lasers for the generation and the detection of acoustic waves. In addition to the spreading of the waveforms due to dispersion, difficulties arise with such materials owing to their high anisotropy and to their small thickness. Their characterization thus requires a proper

TABLE III. Stiffness coefficients and 90% confidence intervals (GPa) measured from experimental signals.

C_{11}^e	C_{22}^e	C_{12}^e	C_{66}^e
8.9±0.1	12.0±0.7	4.9±0.2	2.3±0.1

understanding of the propagation in such a medium and the implementation of an adequate model in order to represent it. Waveforms are then calculated and a well suited signal processing is developed that permits us to calculate the data required by the numerical scheme for the identification. Despite dispersion and echoes overlapping, the calculated group velocity data allow us to recover the stiffness coefficients of the modeled composite material with good reliability. Since the identification process proves to apply properly well despite anisotropy and viscosity, it is used to characterize an actual composite material. Stiffness coefficients are measured and the uncertainty in results is quantified. Involved systematic errors are also discussed. The reliability of most coefficients' identification is sufficient to think about applications of the method in the field of material behavior investigation. For instance, the noncontact characterization of the composite material can be performed to analyze the stiffness properties at elevated temperature. The anisotropic damage[37] induced by temperature in such a material intended to be used in hot structures can then be studied in a phenomenological manner.

ACKNOWLEDGMENT

This work was partially supported by DRET through Grant No. 962532A.

[1] C. B. Scruby, R. J. Dewhurst, D. A. Hutchins, and S. B. Palmer, ''Quantitative studies of thermally generated elastic waves in laser-irradiated metals,'' J. Appl. Phys. **51**, 6210–6216 (1980).
[2] C. B. Scruby and L. E. Drain, *Laser Ultrasonics Techniques and Applications* (Hilger, Bristol, 1990).
[3] R. M. White, ''Generation of elastic waves by transient surface heating,'' J. Appl. Phys. **34**(12), 3559–3567 (1963).
[4] D. A. Hutchins, *Ultrasonic Generation by Pulsed Laser*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1988).
[5] B. Audoin, C. Bescond, and M. Deschamps, ''Measurement of stiffness coefficients of anisotropic materials from point-like generation and detection of acoustic waves,'' J. Appl. Phys. **80**(7), 3760–3771 (1996).
[6] A. G. Every and K. Y. Kim, ''Time domain dynamic response functions of elastically anisotropic solids,'' J. Acoust. Soc. Am. **95**, 2505–2516 (1994).
[7] M. Dubois, F. Enguehard, L. Bertrand, M. Choquet, and J. P. Monchalin, ''Modelling of laser thermoelastic generation of ultrasound in an orthotropic medium,'' Appl. Phys. Lett. **64**, 554–556 (1994).
[8] A. Mourad, M. Deschamps, and B. Castagnède, ''Acoustic waves generated by a transient line source in an anisotropic half space,'' Acust. Acta Acust. **82**, 839–851 (1996).
[9] A. K. Mal and S. Lih, ''Elastodynamic Response of a Unidirectional Composite to Concentrated Surface Loads: Part I,'' Trans. ASME **59**, 878–892 (1992).
[10] R. L. Weaver, W. Sachse, and K. Y. Kim, ''Transient elastic waves in a transversely isotropic plate,'' J. Appl. Mech. **63**, 338–346 (1996).
[11] K. Y. Kim, R. Sribar, and W. Sachse, ''Analytical and optimization procedures for determination of all elastic constants of anisotropic solids from group velocity data measured in symmetry planes,'' J. Appl. Phys. **77**, 5589–5600 (1995).
[12] A. G. Every and W. Sachse, ''Determination of the elastic constants of an anisotropic solid from acoustic-wave group-velocity measurements,'' Phys. Rev. B **42**, 8196–8205 (1990).
[13] K. Y. Kim, W. Sachse, and A. G. Every, ''On the determination of sound speeds in cubic crystals and isotropic media using a broadband ultrasonic point-source/point-receiver method,'' J. Acoust. Soc. Am. **93**, 1393–1406 (1993).
[14] S. Mallat and S. Zhong, ''Characterization of signals from multiscale edges,'' IEEE Trans. Pattern. Anal. Mach. Intell. **14**(7), 710–732 (1992).
[15] K. Aki and G. R. Richards, *Quantitative Seismology* (Freeman, San Francisco, 1980).
[16] D. Royer and E. Dieulesaint, ''Optical detection of sub-Angstrom transient mechanical displacement,'' Proc.-IEEE Ultrason. Symp. **XX**, 527–530 (1986).
[17] B. A. Auld, *Acoustic Fields and Waves in Solids* (Krierger, Malabar, FL, 1990).
[18] M. F. Markham, ''Measurement of the elastic constants of fibre composites by ultrasonics,'' Composites **1**, 145–149 (1970).
[19] A. G. Every and W. Sachse, ''Imaging of laser generated waves in silicon,'' Phys. Rev. B **44**, 6689–6699 (1991).
[20] L. Cagniard, *Réflexion et réfraction des ondes séismiques progressives* (Gauthiers-Villars, Paris, 1939).
[21] R. L. Weaver, W. Sachse, and L. Niu, ''Transient ultrasonic waves in a viscoelastic plate: Theory,'' J. Acoust. Soc. Am. **85**, 2255–2261 (1989).
[22] A. T. De Hoop, ''A modification of Cagniard's method for solving seismic pulse problem,'' Appl. Sci. Res., Sect. B **8**, 349–356 (1960).
[23] C. L. Pekeris, ''The seismic surface pulse,'' Geophysics **41**, 469 (1955).
[24] A. Cand, J. P. Monchalin, and X. Jia, ''Detection of in-plane and out-of-plane ultrasonic displacements by a two channel confocal Fabry-Perot interferometer,'' Appl. Phys. Lett. **64**, 414–416 (1994).
[25] M. Deschamps and B. Hosten, ''Viscoelasticity influence on frequency dependence of the ultrasonic transmission through plates of composite materials,'' Rev. Prog. Quant. Nondestr. Eval. **11**, 201–208 (1992).
[26] P. Guillemain and R. Kronland-Martinet, ''Characterization of acoustic signals through continuous linear time-frequency representations,'' Proc. IEEE **84**(4), 561–585 (1996).
[27] J. J. Benedetto and M. W. Frazier, *Wavelets: Mathematics and Applications*, edited by S. G. Krantz (CCR, London, 1994).
[28] J. Ville, ''Théorie et application de la notion de signal analytique,'' Câbles transm. **2**, 61–74 (1948).
[29] A. Grossmann, M. Holschneider, R. Kronland-Martinet, and J. Morlet, ''Detection of abrupt changes in sound signals with the help of wavelet transforms,'' in *Inverse Problems*, edited by P. C. Sabatier (Academic, New York, 1987), pp. 289–306.
[30] K. Y. Kim, K. C. Bretz, A. G. Every, and W. Sachse, ''Ultrasonic imaging of the group velocity surface about the cubic axis of silicon,'' J. Appl. Phys. **79**, 1857–1863 (1996).
[31] M. Deschamps and C. Bescond, ''Numerical method to recover the elastic constants from ultrasound group velocities,'' Ultrasonics **33**, 205–211 (1995).
[32] B. Audoin, S. Baste, and B. Castagnède, ''Estimation de l'intervalle de confiance des contantes d'élasticité identifiées à partir des vitesses de propagation ultrasonores,'' C. R. Acad. Sci., Ser. II: Mec., Phys., Chim., Sci. Terre Univers **312**, 679–686 (1991).
[33] B. F. Pouet and N. J. P. Rasolofosaon, ''Measurement of broadband intrinsic ultrasonic attenuation and dispersion in solids with laser techniques,'' J. Acoust. Soc. Am. **93**, 1286–1292 (1993).
[34] B. Audoin and C. Bescond, ''Measurement by LASER-Generated Ultrasound of Four Stiffness Coefficient of an Anisotropic Material at Elevated Temperatures,'' J. Nondestruct. Eval. **16**(2), 91–100 (1997).
[35] A. G. Every, W. Sachse, K. Y. Kim, and M. O. Thompson, ''Phonon Focusing and Mode-Conversion Effects in Silicon at Ultrasonic Frequencies,'' Phys. Rev. Lett. **65**(12), 1446–1449 (1990).
[36] H. J. Maris, ''Enhancement of heat pulse in crystal due to elastic anisotropy,'' J. Acoust. Soc. Am. **50**, 812–818 (1971).
[37] S. Baste and B. Audoin, ''On internal variable in anisotropic damage,'' Eur. J. Mech. A/Solids **10**, 587–606 (1991).

Effects of higher-order modes and harmonics in single-bubble sonoluminescence

Fred B. Seeley
Physics Department, University of Alabama Huntsville, Huntsville, Alabama 35899

(Received 21 January 1998; revised 20 October 1998; accepted 21 December 1998)

This paper presents the results of a series of experiments that was performed to study the higher-order modes and harmonics that exist within water-filled resonators during single-bubble sonoluminescence [SBSL]. These experiments demonstrated that: (1) the spatial position and the phase of the light flashes of an SBSL bubble relative to the fundamental sound field could be significantly changed by adding higher harmonics, (2) SBSL could be produced in spherical flasks at eigenfrequencies of the higher-order modes and higher harmonics of the spherical Bessel functions, j_1, and j_2, thus producing noncentered flashing bubbles, and (3) it is possible to produce the SBSL effect at eigenfrequencies at which the eigenmodes are degenerate. These experiments show that a rich, complex mode structure evidently exists within the water-filled resonator. The externally generated higher modes and higher harmonics influence the position, spatial stability, and phase of the light-flash characteristic of single-bubble sonoluminescence. These experiments suggest that any higher-order modes and harmonics, self-generated by the collapsing bubble, may spatially shift their position within the resonator and thus may play a significant role in destabilizing the bubble's motion. © *1999 Acoustical Society of America.* [S0001-4966(99)00204-0]

PACS numbers: 43.35.Ei [HEB]

INTRODUCTION

Whenever single-bubble sonoluminescence [SBSL] is produced in a water-filled resonator, acoustic oscillations can be detected at the wall of the resonator using a small piezoelectric transducer [PZT] attached to the outer wall.[1,2] Periodic pressure impulses can be measured at locations near the bubble using a needle hydrophone[3,4] or focused transducers.[5] The largest, initial pulse appears to originate at the instant the bubble reaches its minimum radius and the light flash occurs.[5] Moreover, the acoustic oscillations observed at the wall of the resonator or the hydrophone output near the bubble have an acoustic spectrum that contains hundreds of harmonics. (See Fig. 1.) This spectrum is consistent with a Fourier spectrum of the narrow, periodic pulse that occurs once per bubble collapse. Both the pressure impulse observed near the bubble and the acoustic emissions are phase-locked to the light flashes.

Previous studies have referred to these pulses and oscillations as ''emissions from the single bubble.''[4,5] A slightly different view was hypothesized and guided the experiments described in this paper: it was assumed that the periodic pressure impulse from the SBSL bubble ''pumps'' harmonic eigenfrequencies in the water-filled resonator. This simple, intuitive model assumes that the pressure impulses from the bubble's collapse cause a ringing in the water-filled resonator itself. This leads to the idea that the hydrophones and wall-mounted PZT respond to the net superposition of all the excited eigenfrequencies in the water. The presence of the highly reflecting resonator wall returns the energy from each collapse, and a periodic wave structure is built up within the water itself over a few hundred to a few thousand cycles of the fundamental, depending upon the Q of the resonator. Of course, only those eigenmodes that were also harmonics of the driving fundamental would be excited. The superposition of all the eigenmodes and eigenfrequencies within the resonator would result in the complex temporal and spatial distribution of acoustic pressure that has been repeatedly observed during SBSL operation.

This paper describes our efforts to excite higher-order modes and higher harmonics of these modes in water-filled resonators typically used in SBSL. It was determined that many eigenmodes can not only be excited, but will support SBSL production. The superposition of degenerate modes and harmonically related modes was experimentally done, also producing SBSL with lighted bubbles. However, testing the idea that a superposition of eigenmodes and harmonics within the water-filled resonator is physically possible was not the only goal. The principal motivation was to investigate whether externally generated high-frequency acoustic pressure oscillations and pressure pulses have any physical effects on the phase, spatial stability, or the luminescent intensity of the sonoluminescing bubble. Moss, working from Lawrence Livermore National Laboratory (LLNL), has suggested and calculated that adding an external pressure spike, properly synchronized with the collapse and at the bubble's location, should increase its brightness.[6] Our preliminary attempts to implement Moss's suggestion were unsuccessful. To understand the processes, the investigation looked into how the water-filled resonator and the sonoluminescing bubble responded as external pressure oscillations and impulses were applied during single-bubble sonoluminescence.

Similar work has already been done in enhancing cavitational effects by adding harmonics. Umemura *et al.* demonstrated the effects of two-frequency forcing of cavitating medical treatment systems in which they used traveling waves and the superposition of the 1st and 2nd harmonics.[7] In these medical studies, cavitation and sonochemical reactions were enhanced by an order of magnitude when the intensities of the 1st and 2nd harmonics (0.5 and 1.0 MHz) were approximately equal and the phase difference maxi-

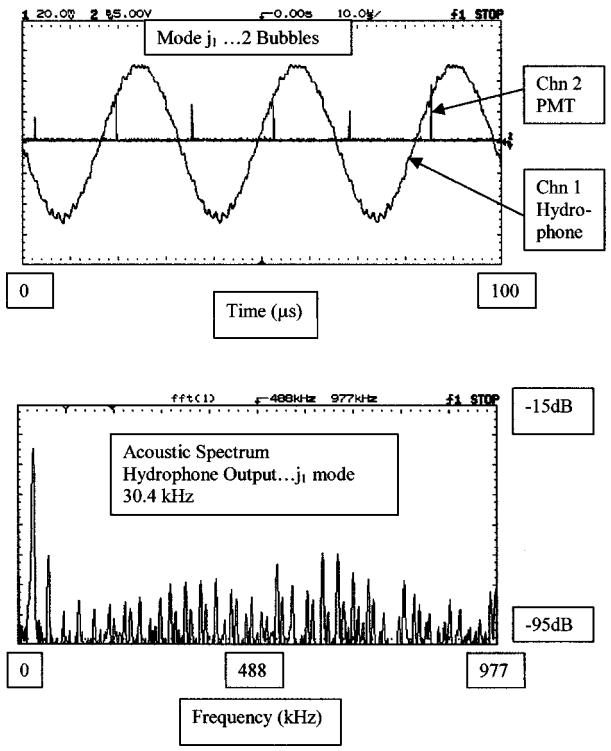


FIG. 1. A typical output showing the timing of the light flash from a pair of SL bubbles, the acoustic driving pressure, and the higher-harmonic oscillations superimposed on the sinusoidal drive of the fundamental. In this example, the resonator is spherical and is being driven at a mode that has two antinodes, with a single bubble at each. Note that the two bubbles flash out of phase. Also shown is the Fourier-transformed hydrophone signal illustrating the periodic spectrum in the frequency domain.

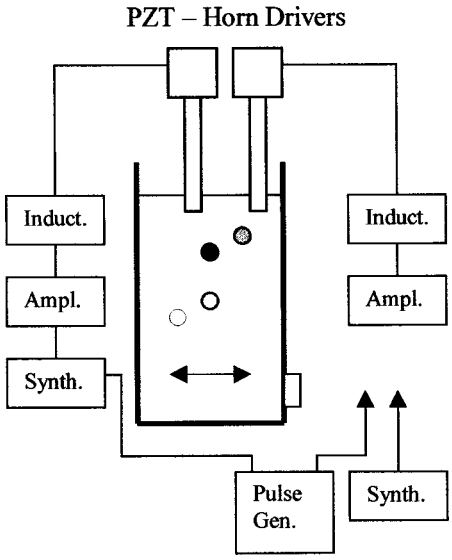


FIG. 2. This figure schematically shows the apparatus used to excite a rectangular cell being driven at two modes, [1,1,3], 29.009 kHz, and [2,2,6], 58.198 kHz in a 5.8×5.8×10.04-cm rectangular box. The second harmonic displaces the flashing bubbles in opposite directions. For a 2nd harmonic acoustic pressure of ~0.1 atm, the bubbles were each displaced approximately 6 mm from their original positions while continuing to sonoluminesce. The shading of the bubbles indicate that the bubbles flash out of phase. Spheres, cylinders, or other containers were excited and produced similar results.

mized the peak rarefaction of the traveling wave. It has also been reported that cavitation can be reduced by phase modulating the exciting signal.[8]

In contrast with these earlier reports, the experiments to be described in this paper involved the responses of single, sonoluminescing bubbles to perturbations to the single frequency and single mode, standing-wave acoustic fields ordinarily used to drive an SBSL, high-Q resonator. As more harmonics of higher-frequency standing waves are externally added, the positive and negative peak pressures no longer remain in the same position within the resonator. Thus, the resulting wave superposition becomes a type of traveling wave. The presence of higher harmonics also increases the acoustic pressure gradients, which in turn increase the forces acting to displace any bubbles present in the fluid.[9] Therefore, generation of both the traveling waves and the larger pressure gradients should allow manipulation of the levitated position of the single sonoluminescing bubble. This would be a result similar to that already well understood and exploited in acoustic levitation experiments in which liquid and solids have been moved and precisely positioned within resonant acoustic chambers.[10,11] Displacement of the SBSL bubble was the observed result.

Although the experiments described in this paper used external excitation, there is a potential connection to SBSL. It is reasonable to assume that any complex standing and traveling waves within the water-filled resonator, regardless of their origin, will have similar effects on the sonoluminescing bubble. If, in fact, the collapsing bubble generates a complex mode structure within the resonator, the bubble's position, phase, and even radial motion could all be influenced in the same way as it is by externally generated modes.

I. OUTLINE OF EXPERIMENTS

In order to study experimentally the potential role played by higher-order harmonics and modes, three major series of tests were performed. These included:

(1) Addition of single harmonics from external oscillators to modify the harmonics self-generated by the collapse of the sonoluminescing bubble. (Some preliminary but unsuccessful attempts also were made to apply pressure impulses as suggested by Moss.)

(2) Production of SBSL in spheres at modes other than $j_0 = \sin(kr)/kr$, specifically, at several harmonics of the spherical Bessel functions,

$$j_1 = \frac{\sin(kr)}{(kr)^2} - \frac{\cos(kr)}{kr},$$

and

$$j_2 = \left[\frac{3}{(kr)^3} - \frac{1}{kr}\right]\sin(kr) - \frac{3}{(kr)^2}\cos(kr).$$

(3) Production of SBSL at the degenerate eigenmodes of box and cubical resonators.

II. GENERAL APPARATUS

Figure 2 schematically illustrates the apparatus used to vibrate several types of water-filled chambers at one of its resonant eigenfrequencies. Several different chamber geom-

etries and sizes were used. The water was excited using a technique that differs from those commonly used in SBSL research.[1,2,4,5] The water-filled resonator was driven by inserting the free end of a vibrating rod 2 to 4 mm into the open surface of the water, making a type of low-power, tunable-frequency sonicating horn. This is not only a very efficient manner of coupling energy into the water but, by moving the position of the rod, it is also a means to control the phase of the modes that are being excited. Both a commercial sonicating horn and locally fabricated horns were used. During the series of experiments when the fundamental and a harmonic were being simultaneously excited, two separate piezoelectric drivers were used. Initially, a specially fabricated horn was used that had two coaxially mounted tube piezoelectric drivers on a single rod. However, it was found to be more effective to use two separate horns so the higher harmonic and modes could be excited in a stable manner. Using tuned circuits, electrical inputs to the horns of generally less than 1 watt were required to excite the water-filled resonators at a range of driving eigenfrequencies from 15 to 95 kHz. The externally driven harmonics were at 2 to 20 times the driving fundamental. It was not necessary to mechanically change the length of the horns to successfully produce the 1.3-atm acoustic pressure needed for SBSL production. This same excitation method was used for spherical, cylindrical, rectangular box-like, and cubical resonators. The spheres, cubes, and rectangular boxes were all made of 2-mm-thick acrylic. The cylinders were made from acrylic tubes with 4-mm walls with a 2-mm acrylic cap on the bottom. The use of such a thin wall made it possible to use the exterior dimensions of the resonators as the dimensions of the free-boundary condition.[12] When substituted into the appropriate solutions of the wave equation for spherical, cylindrical, and rectangular resonators, the predicted and measured eigenfrequencies typically agreed to within 2 percent. Supporting instruments used but not illustrated in Fig. 2 include several general-purpose oscilloscopes, an HP 54603B digital oscilloscope *w*/fft plug-in, and a photomultiplier tube. Wall-mounted PZT monitors and Dynasen PIN, CA1136 pressure detectors were both used to measure the acoustic responses at the wall and within the water.

III. SINGLE HIGHER-HARMONIC EXPERIMENTS

A. Objective and approach

In this first series of tests, a single additional high-frequency excitation was externally added to the water-filled resonator during SBSL operation. The objective of this series of experiments was to determine what effect adding single harmonics to the existing standing-pressure wave structure would have on the sonoluminescing bubble; specifically, could the light output be increased or the bubble position be made more stable? Initially, the externally applied harmonic was independently generated using a second frequency synthesizer which had ten times the resolution of the generator used to drive the fundamental. Normally, no attempt was made to control the phase and lock the frequency of the added signal to the harmonics of the SBSL generator. (Later, the apparatus was modified to allow controlling the phase of the 2nd and 3rd harmonics relative to the fundamental.) When used as an independent electrical signal, the second piezoelectric transducer (PZT) was driven at frequencies closely set at harmonics ranging from two to 20 times the fundamental frequency. The driving amplitudes for this second PZT were adjusted to increase the amplitude of the high-frequency components seen during normal SBSL operations. These experiments were performed in spherical, rectangular, and cylindrical cells. It was found that the single high-frequency component could be added to the resonator with either a second horn or by using a PZT attached directly to the resonator wall to produce similar effects.

B. Results

In all cases, the photomultiplier did not detect an increase in the brightness of the bubble as a result of adding a single, higher harmonic. However, what did occur was a precise perturbation in the position of the flashing bubble. When the frequency of the added higher harmonic was adjusted to produce a low-frequency beat with one of the bubble's harmonics, the bubble would travel a periodic, though sometimes complex, orbit about the equilibrium position. The frequency of this motion equaled the beat frequency that was chosen. This translation in the bubble's position occurred while the bubble continued to luminesce. Amplitudes of the motion were as large as 6 mm; i.e., the peak-to-peak displacement of the lighted bubble was 1.2 cm. Figure 2 shows a particular result for a rectangular cell excited at two eigenfrequencies corresponding to the eigenmodes [1,1,3] and [2,2,6], where the modes are described by the integral values in $[n_x, n_y, n_z]$ in the solution to the wave equation in the rectangular resonator. In the example illustrated in Fig. 2, a pressure amplitude of approximately 0.1 atm in the [2,2,6] mode pulled the lighted bubble approximately 6 mm from the antinode of the fundamental. In other cases, using different types of resonators and drive techniques, linear, elliptical, and more complex orbits were observed moving at orbital frequencies from near zero to over 20 Hz. When two bubbles were lighted in say, the [1,1,2] mode of a rectangular cell, the bubbles would move 180 deg out of phase. If the motion were primarily vertical, one bubble would translate upward as the other would move down. When the motion was elliptical, one bubble would move clockwise as the other moved counterclockwise. These kinds of orbital motions were produced in spherical, rectangular, cubical, and cylindrical cells. A driving PZT on the side wall, the special, two-coaxial PZT horn, or two separate horns could all produce the effects just described.

The added harmonic experiment was subsequently reperformed using apparatus that allowed controlling the relative phases of either the 2nd or 3rd harmonics relative to the fundamental drive. Using the same box resonator shown in Fig. 2, SBSL was produced in the [1,1,1] mode at approximately 19.6 kHz. The [2,2,2] mode was excited at 39.2 kHz. This [2,2,2] eigenmode produced the maximum bubble displacements when its phase reinforced the negative and positive cycles of the fundamental standing wave. The bubble would be pulled into that volume of the resonator where the negative (and positive) pressure of the [1,1,1] mode was be-

ing reinforced by the added [2,2,2] mode, i.e., in direction of the maximum—grad P. A typical maximum bubble displacement for this case was 1.5 mm when the acoustic pressure of [2,2,2] was ~0.2 atm. By manually adjusting the relative phase, the bubble could be displaced and made to travel in orbits similar to those observed in the experiments without phase control.

The phase-controlled experiment with an odd harmonic at [3,3,3] produced similar displacements, but differed with respect to the relative phasing. Adjusting the relative phase of the 3rd harmonic to reinforce the maximum and minimum acoustic pressures of the fundamental produced a minimum displacement of the bubble. A further phase shift of 90 deg in the 3rd harmonic would then produce the maximum bubble displacement. A 1.5-mm displacement could be produced with ~0.03-atm acoustic pressure for mode [3,3,3] at 58.8 kHz. It must be emphasized that the measurements and observations were made while the bubble continued to sonoluminesce, flashing at the repetition rate of the fundamental frequency. If the bubble were displaced too far from the antinode of [1,1,1], the bubble would cease visibly flashing. However, the dark bubble could still be observed using scattered light. When the 2nd or 3rd harmonic amplitude or phase was adjusted to allow the bubble to move closer to the fundamental's antinode, it would resume visibly flashing.

Although no increase in the brightness of the bubble was ever observed using a photomultiplier, the light flash was seen to be affected in two interesting ways when a higher harmonic was externally added. First, the phase of the flash relative to the fundamental drive could be shifted by as much as 5 μs during acoustic periods of approximately 50 μs. This is of the same order as the shifts in the flash phase that occur when the fundamental driving frequency is tuned across the resonance. Second, the light flash could be interrupted and then turned back on again. This would regularly occur at the same points in the orbit of the bubble. Thus, adding energy at just a single harmonic could move the sonoluminescing bubble and also perturb the phase of the light flash relative to the fundamental drive. These effects are indirect evidence that the position and timing of the light flash also may be influenced by the series of higher harmonics that the bubble itself produces.

C. Preliminary pulse tests

As mentioned in the introduction, Moss from LLNL suggested that adding a 3-atm pressure spike during the collapse phase should increase the brightness and temperature of a sonoluminescing bubble during the <50 picosecond flash. Motivated by his suggestion, we did some preliminary work to test this idea, but were unsuccessful in increasing the brightness of the sonoluminescent flash. Basically, the experiments were an extension of the single-added harmonic tests; to add a large number of harmonics by equivalently adding a narrow, periodic pulse of acoustic energy. During these pulse tests, the repetition frequency of the pulse was locked to that of the sinusoidal fundamental. The phase of a narrow electrical pulse with respect to the light flash was carefully adjusted over a full 360 deg. Electrical pulses of 20 to 60 v peak-to-peak were applied to 1-cm-diameter PZTs attached to the side walls of spherical and rectangular resonators.

Results similar to that of the single harmonic experiments were obtained; i.e., small, 1- to 5-mm, perturbations in the bubble's position were produced. No effect on the brightness of the bubble was detected using the photomultiplier tube. However, as the phase of the pulsed PZT was varied over a full 360 deg, the perturbation in position could be varied. Electrical inputs to the PZT that were positive going, negative going, and bipolar were all tried. These simple tests were terminated when it was realized that electrically pulsing a single PZT did not guarantee that a narrow, pressure pulse was being produced at the bubble's position near the antinode(s) of the fundamental within the water-filled resonator.

IV. HIGHER-ORDER MODES AND HARMONICS WITHIN A SPHERICAL RESONATOR

A. Objective

The second series of tests used higher-order modes in a water-filled, acrylic sphere. The objective of these tests was to determine if single-bubble sonoluminescence (SBSL) could be produced in a spherical resonator by exciting acoustic pressure modes other than $j_0 = \sin(kr)/kr$. Because the spherical Bessel functions j_1 through j_n all are zero at the center of a sphere, if eigenfrequencies of these higher-order modes could be excited, then it might be possible to produce SBSL at antinodes well removed from the sphere's center. The secondary objective was to determine if any differences existed in the brightness or acoustic spectra when SBSL was produced using these higher-order modes.

B. Method and apparatus for higher-order modes in a sphere

The commercial insonicator horn was inserted into the free-water surface to drive spherical acrylic resonators at the desired eigenfrequencies. The circuit was electrically tuned by adjusting the inductors to make the driver circuit resonant at the selected eigenfrequencies from 19 to 95 kHz. A 2-mm-thick acrylic spherical shell, outer radius $a = 3.50$ cm, was used. A 3/4-in.-diameter hole was drilled in the top to insert the 1/2-in.-diameter sonicator horn. PZT monitors were epoxied to the bottom and to the upper hemispheres of the shell. The sphere itself was mounted on a three-point kinematic stand in contact with 1/2-in. foam boards. The bubbles were seeded using a 1-Ω heater wire fed in through the sphere opening. The wire was repositioned as needed to produce pairs of bubbles.

C. Results

It was possible to excite many modes and many eigenfrequencies of the higher-order modes, j_1 and j_2. Also, it was discovered that sonoluminescence could be produced at several eigenfrequencies of both j_1 and j_2. By assuming that the acrylic wall acted like a free boundary, the calculated and measured eigenfrequencies agreed to better than 2 percent. (This simplified tuning to the correct eigenfrequency.) Table

TABLE I. Eigenfrequencies of spherical Bessel functions j_1 and j_2 in a 35-mm-radius sphere. Calculated and measured eigenfrequecies agree to better than 0.5%. Displayed data was obtained in a single SBSL session, in less than 8 h using a single filling of the sphere with degassed water. During other sessions, SBSL has been produced at the two eigenfrequencies, 73.9 and 61.7 kHz.

	Eigenfrequencies [kHz]		
Zeros (j_0)	Calculated	Measured	SBSL
π	21.310	21.270	yes; 1 centered
2π	42.619	42.400	yes; 1 centered
3π	63.925	63.319	yes; 3 bubbles
4π	85.240	85.801	yes; 1 centered
	Eigenfrequencies [kHz]		
Zeros (j_1)	Calculated	Measured	SBSL
4.493	30.477	30.590	yes; 2 out of phase
7.724	52.395	52.537	yes; 3
10.904	73.964	73.910	no; resonance only
14.066	95.421	95.362	yes; 2 out of phase
	Eigenfrequencies [kHz]		
Zeros (j_2)	Calculated	Measured	SBSL
5.754	39.098	39.054	yes; 2 in phase
9.095	61.693	61.453	no; resonance only

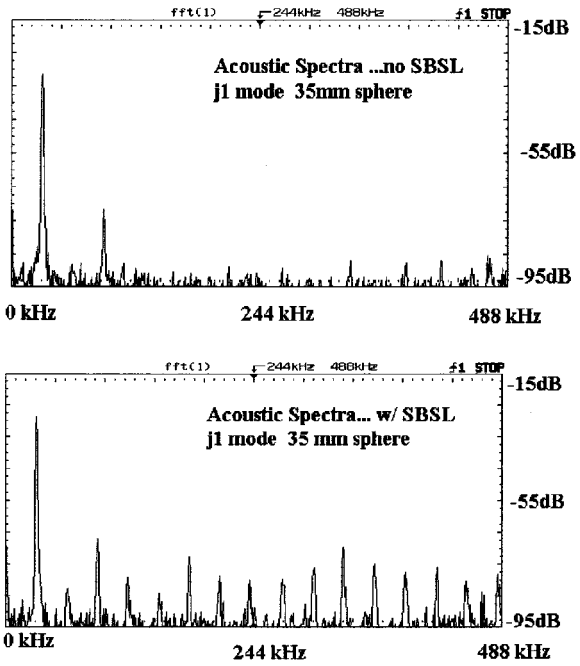


FIG. 3. The acoustic spectrum of the output from a PZT pressure detector extends beyond the 488 kHz illustrated in this recording to at least 2.5 MHz. This spectrum was recorded while two bubbles were flashing out of phase at 30.72 kHz, the first eigenfrequency of the j_1 mode in a 35-mm-radius sphere. The tip of the detector was ~4 mm from one of the SL bubbles. The second spectrum is for the same driving excitation but with no bubble(s).

I summarizes typical observations for the 35-mm-radius sphere. The lighted bubbles were fixed near the pressure antinodes, i.e., not at the center of the sphere. In the case of eigenfrequencies of j_1, an odd function, two bubbles could be trapped and lighted, flashing out of phase as expected. Eigenfrequencies of j_2 would also trap two or more bubbles which would flash in phase, since j_2 is an even function. The bubbles excited in the j_1 and j_2 modes exhibited the same transitions in the motion and the same narrow, 20%–30% range in acoustic pressure during luminescence that one ordinarily sees in the j_0 mode. Precise measurements of the position of the lighted bubbles were difficult to make. Nevertheless, we estimated that to within 2 to 4 mm, the lighted bubbles were at the pressure antinodes. Measurement of the acoustic spectra of the j_1 and j_2 modes during SBSL show a harmonic structure, up to at least 2 MHz, very similar to that observed when eigenfrequencies of mode j_0 are used. Figure 3 shows the acoustic spectrum of the hydrophone output in the j_1 mode both with and without the paired bubbles at the two antinodes in the sphere.

D. Significance of spherical tests

This experiment showed that the spherical Bessel functions j_1 and j_2 are both physically possible modes within a spherical resonator. Moreover, bubbles can be captured, levitated, and SBSL produced at the pressure antinodes of these modes. This brings up the interesting point that degenerate modes can be produced within a water-filled sphere. The degeneracies can arise in two ways: from the analytic solutions and in an approximate form. First, the solutions to the spherical wave equations include spherical Bessel functions of order l which are $2l+1$ degenerate. For example, when $l=1$, there are three different angular distributions of the acoustic pressure which are valid solutions to the wave equation. These each have the same eigenfrequency, which is determined by the radial function, $j_1(kr)$. Therefore, even though mathematically degenerate, it is physically possible to excite these modes in a stable manner and to produce SBSL.

Another approximate-type degeneracy can also occur. Zeros of all the even-valued spherical Bessel functions, j_0, j_2, j_4, etc., asymptotically approach the same value at high mode numbers. Because the Q's of these acoustic resonators are on the order of 500 to 5000, this approximate degeneracy may become effective beginning approximately at the 5th harmonic of j_0. Because the 5th harmonic of j_0 is approximately equal to the 4th eigenfrequency of j_2, energy can be coupled from one mode to the other. Any other even spherical Bessel function can also approximate an eigenfrequency of a higher harmonic. Thus, any excitation, external or that self-generated by the imploding bubble, could excite these approximately degenerate modes when the harmonic numbers become high enough. The collapsing bubble generates an acoustic spectrum of harmonics which extends well into the range where this approximate degeneracy is possible in the sphere. Obviously, the same conditions for mode degeneracy exist within cylindrical resonators where the radial solutions to the wave equation take the form of ordinary Bessel functions. The question then is, are these approximately degenerate modes inherently unstable, causing a low-frequency modulation of the spatial pressure distribution as energy is being coupled from one mode to another? Further, could the degenerate modes be excited and be stable enough to produce SBSL?

V. DEGENERATE MODE TESTS

The third series of tests further explored experimentally the conditions of stability of degenerate modes. To do this, degenerate acoustic modes were produced in cubical and parallelepiped water-filled resonators. The specific objective was to determine if the degenerate modes could ever be excited to be stable and, specifically, if single-bubble sonoluminescence (SBSL) could be produced. These tests were limited to degeneracies in the cubical and box resonators.

The same general apparatus illustrated in Fig. 2 was used, filling the resonator cell with degassed water to form either a cube or parallelepiped of the desired depth. The cube was approximately 5.8 cm on a side and at the mode [1,1,2] is threefold degenerate with an eigenfrequency of 31.6 kHz. Various placements of the sonicator horn were used to determine if a way could be found to excite all three modes simultaneously and be stable. The tests were also performed on the next degenerate eigenfrequency of the cube, [1,2,2]. A box $5.8\times5.8\times10.04$ cm is degenerate at the three modes; [1,1,5], [1,3,1], and [3,1,1]. Tests of this degeneracy were also conducted.

It was found that SBSL could be produced using the degenerate modes in the cubical and box resonators that were tested. An approximately equal excitation was produced in all three modes at each eigenfrequency by positioning the insonicating horn centered in one of the four quadrants of the top face. This resulted in a pair of pressure antinodes symmetrically located within the resonator volume along one of the diagonals of the cube. Figure 4 shows the calculated superposition of three cubical modes, [1,1,2], [1,2,1], and [2,1,1], assuming that they are of equal pressure amplitude and all temporally in phase. In this case, the maximum pressure antinodes are 180 deg out of phase. This positioning and phase was confirmed by the experimental results. Pairs of bubbles produced flashes that also were out of phase and located near the predicted positions along the diagonals of the cubes, at $X=Y=Z=1/3$ and 2/3. Figure 5 illustrates these results. By relocating the position of the insonicating horn into different quadrants of the top face, the relative phases of the three modes could be altered and the bubble pairs shifted in position to lie along other diagonals of the cube.

Preferential excitation of one of the degenerate modes was unstable and SBSL could not be achieved. In the [1,1,2] example, whenever the sonicator horn was centered in the top face of the cube, the water surface would oscillate violently at 1 to 10 Hz, producing surface waves. The PZT monitor also would indicate 50% to 100% variations in its indication of the wall displacement. Evidently, this low-frequency instability was caused by preferential excitation of mode [1,1,2], which would then couple to the other modes. A fuller explanation of this low-frequency instability with surface oscillation may be similar to the production of fluid rotation by out-phase orthogonal waves reported by Biswas *et al.*[13] For example, the mathematical superposition of modes [1,1,2] and [1,2,1] that are 90 deg out of phase predicts that the pressure maximum and minimum would rotate about the central axis of the cube at a frequency equal to that of the eigenfrequency. Instrumentation was not available to attempt detect any rotation of the pressure maxima or of the fluid itself. However, positioning the horn tip just below the water surface in any one of the four top quadrants resulted in stable excitation, i.e., this position of the horn eliminated the low-frequency oscillation and the surface waves, and also resulted in a stable output signal from the PZT monitor. With this stable excitation of the degenerate modes, SBSL operation was then possible. During stable SBSL operation with equal mode excitation, the brightness of the bubble and transitions in motion were indistinguishable from a bubble trapped in a nondegenerate mode such as [1,1,1].

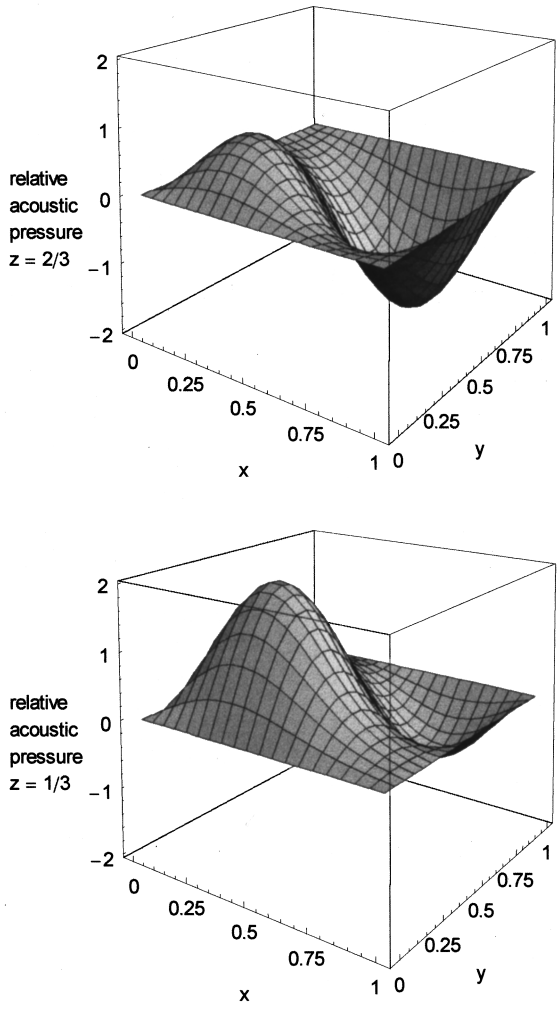


FIG. 4. Plot of the calculated superposition of the pressure amplitudes of three degenerate eigenmodes, [1,1,2], [1,2,1], and [2,1,1] in a cube 1.0 unit on a side at two different vertical cross sections, $z=1/3$ and 2/3. The calculation assumes that the pressure amplitudes of each mode are all equal and in phase, resulting in two pressure antinodes that are perfectly antisymmetric and lie along the diagonal of the cube at $x=y=z=1/3$ and 2/3. The X and Y axes are the coordinates of the cube, which vary from 0 to 1.0. The vertical axis is proportional to the acoustic pressure.

The prediction for a cube at [1,2,2] was for similarly located bubbles within the cube at the antinodes also positioned along the diagonal, but this time oscillating and flashing in phase. This also was the result observed. Out-of-phase flashes were predicted for the taller $5.8\times5.8\times10.04$-cm box operating at the three modes, [1,1,5], [1,3,1], and [3,1,1]. Only pairs of bubbles were lighted. However, by adjusting the drive amplitude, bubbles could be trapped and lighted at alternate paired locations, confirming the predictions.

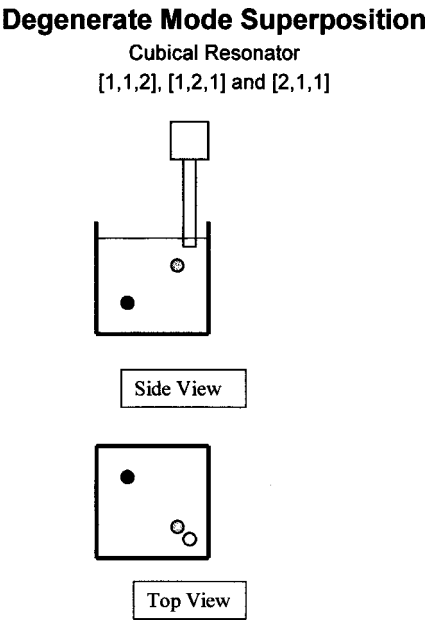


FIG. 5. Illustration of the diagonally placed pair of flashing bubbles within the cube excited at the degenerate mode [1,1,2]. The shading in this figure indicates that the bubbles flash out of phase, which is experimentally verified by observation with a photomultiplier. The position of the horn tip determines along which diagonal the bubbles will be captured.

VI. SUMMARY AND OVERALL CONCLUSIONS

The three series of tests studied the rich, complex mode structure that can be physically excited within water-filled resonators of the type used to produce single-bubble sonoluminescence. The common result from these experiments is that the net acoustic pressure within the water-filled resonator may be considered as a superposition of all the harmonics and all possible modes of each harmonic. It is well known that box and cubical resonators have higher-order modes that can be excited and used to produce sonoluminescence. The spherical mode tests showed that higher-order modes that mathematically exist as solutions to the spherical wave equation also could be excited and used to produce SBSL. By capturing and lighting a bubble, it was conclusively shown that the higher-order spherical Bessel function modes, even though mathematically degenerate, were stable. Moreover, these modes had pressure antinodes well removed from the center of a sphere.

The tests of the degenerate modes in the cubes and boxes showed that stable SBSL could be achieved under special conditions, namely approximately equal excitation of all modes. These degeneracy tests also showed that the lighted bubbles would be located at the pressure antinodes that are the superpositions of the three different modes. It was also noted that the degenerate modes were unstable if one mode were externally excited more than another. Under this preferential excitation, the energy from one mode would couple to the other resulting in a low-frequency vibration of the water-filled box or cubical resonator. By contrast, because of the symmetry of the spherical resonator, this low-frequency vibration oscillation was not observed while exciting the degenerate j_1 and j_2 modes with the external horn.

Finally, it was shown that when a single harmonic excitation was added to that of the fundamental drive frequency, the SBSL bubble was spatially displaced. This effect is dependent upon the mode and phase as well as the acoustic pressure gradient of the added harmonic. Because the differences between adjacent eigenfrequencies of any resonator become smaller with increasing harmonic number, the probability of exciting multiple modes increases at higher harmonics. In the added-harmonic tests, external sources of energy were supplied via a second electrical oscillator and PZT, thus producing a stable superposition of modes and a stable spatial perturbation in the position of the bubble. However, when the resonator is being externally excited by just the fundamental drive, all the harmonics must originate from the motion of the bubble. If there were any shift in the bubble's position caused by its self-generated harmonics and modes, then the shift would likely be unstable because one mode is preferentially excited. Therefore, understanding how the pressure impulse from the SBSL bubble excites each harmonic and mode may yield additional insight about the instability of the single, sonoluminescing bubble.

ACKNOWLEDGMENTS

Support was received from the UAH Research Institute and from the Center for Microgravity and Materials Research at UAH. Acknowledgment is made for the excellent technical advice given by the SBSL research groups at the University of Mississippi, UCLA, LLNL, and at the University of Washington.

[1] D. F. Gaitan and L. A. Crum, ''Observation of sonoluminescence from a single stable bubble in a water/glycerin mixture,'' in *Frontiers of Nonlinear Acoustics: 12th International Symposium on Non-Linear Acoustics*, edited by M. Hamilton and D. Blackstock (Elsevier, New York, 1990), pp. 459–463.

[2] B. P. Barber, ''Synchronous picosecond sonoluminescence,'' UCLA Ph.D. thesis (1992).

[3] K. R Weninger, B. P. Barber, and S. Putterman, ''Pulsed Mie scattering measurements of a sonoluminescing bubble,'' Phys. Rev. Lett. **78**, 1799–1802 (1997).

[4] T. J. Matula, S. M. Cordry, R. R. Roy, and L. A. Crum, ''Bjerknes force and bubble levitation under single-bubble sonoluminescence conditions,'' J. Acoust. Soc. Am. **102**, 1522–1527 (1997); S. M. Cordry, ''Bjerknes forces and temperature effects in single-bubble sonoluminescence,'' Ph.D. thesis, University of Miss., 1995.

[5] T. Matula, I. Hallaj, R. Cleveland, L. Crum, W. Moss, and R. Roy, ''The acoustic emissions from single-bubble sonoluminescence,'' J. Acoust. Soc. Am. **103**, 1377–1382 (1998).

[6] W. C. Moss, D. B. Clark, J. W. White, and D. A. Young, ''Sonoluminescence and the prospects for table-top thermonuclear fusion,'' Phys. Lett. A **211**, 69–74 (1996).

[7] S. Umemura, K. Kawabata, and K. Sasaki, ''*In vitro* and *in vivo* enhancement of sonodynamically active cavitation by second-harmonic superimposition,'' J. Acoust. Soc. Am. **101**, 569–577 (1997).

[8] J. Y. Chapelon, F. Dupenloup, H. Cohen, and P. Lenz, ''Reduction of cavitation using psedorandom signals,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 623–625 (1996).

[9] A. Eller, ''Force on a bubble in a standing acoustic wave,'' J. Acoust. Soc. Am. **43**, 170–171 (1967).

[10] E. Trinh, J. Robey, N. Jacobi, and T. Wang, ''Dual-temperature acoustic levitation and sample transport apparatus,'' J. Acoust. Soc. Am. **79**, 604–612 (1986).

[11] S. Min, G. Holt, and R. Apfel, ''Simulation of drop dynamics in an acoustic positioning chamber,'' J. Acoust. Soc. Am. **91**, 3157–65 (1992).

[12] F. B. Seeley and C. K. Joens, ''Synchronous sonoluminescence in acrylic resonant chambers,'' Am. J. Phys. **66**, 259–260 (1998).

[13] A. Biswas, E. W. Leung, and E. H. Trinh, ''Rotation of ultrasonically levitated drops,'' J. Acoust. Soc. Am. **90**, 1502–07 (1991). Phys. Rev. Lett. **69**, 1182 (1992).

Fast numerical scheme of computing acoustic pressure fields for planar circular ultrasound transducers

Chihng-Tsung Liauh
Department of Mechanical Engineering, Kung-shan Institute of Technology, Tainan

Win-Li Lin[a)]
Institute of Biomedical Engineering, College of Medicine, National Taiwan University, No. 1, Sec. 1, Jen-Ai Road, Taipei, Taiwan

(Received 27 April 1998; revised 15 December 1998; accepted 26 December 1998)

The velocity potential field and thus the acoustic-pressure field of planar circular ultrasound transducers are theoretically derived. By applying the Hankel transform, the solution of the wave equation induced by the transducer can be expressed as a single integral in which both the wavelength of the emitted ultrasound and the radius of the transducer become mere parameters. Two-dimensional surface and contour plots of acoustic pressure fields for planar circular transducers with different values of a/λ, the ratio of transducer radius to wavelength, are compared with previously published plots calculated by other methods. The comparison reveals that the computational time in obtaining accurate calculations is dramatically reduced for high frequency and/or large transducers. © *1999 Acoustical Society of America.* [S0001-4966(99)02404-2]

PACS numbers: 43.35.Yb, 43.88.Ar [HEB]

INTRODUCTION

According to the theory of acoustic radiators presented by Kinsler and Frey,[1] the velocity potential (ϕ) of the transducer could be calculated by the Rayleigh–Sommerfeld integral

$$\phi = \frac{1}{2\pi} \int \int_A \frac{u_0 e^{-j\kappa s}}{s} dA,$$

where $j = \sqrt{-1}$, u_0 is the velocity normal to the transducer surface, k is the wave number of the propagation wave, and s is the distance between the source point dA and the field point p as shown in Fig. 1. Zemanek[2] presented the two-dimensional surface and contour plots of the pressure distribution obtained by numerical integration of the double integral. Zemanek also demonstrated that obtaining accurate pressure distributions using a two-dimensional numerical integration over the entire transducer surface was quite time-consuming. The reason was that in order to be considered as point sources, the grid size for the computations had to be sufficiently small when compared with the wavelength.

However, rapid numerical approximations have been proposed to reduce the computational time in calculating the acoustic velocity potential distribution generated by a transducer. Lockwood and Willette[3] contended that the pressure variation at any point in the field of a given baffled piston could be computed by evaluating the driving frequency component of the Fourier transform of the impulse response. This approach required only a single numerical integration, while in comparison Zemanek's double integration required extra steps by more than a factor of a/λ. Roemer *et al.*[4] proposed that if the normal velocity remained constant over the entire transducer surface, the velocity potential at a field point p could be expressed as

$$\phi_p = \frac{u_0}{2\pi} \sum_{m=1}^{M} \frac{\Delta A_{m,p}}{s_{m,p}} e^{-jks_{m,p}},$$

where $\Delta A_{m,p}$ is the element area over which the wavelet with radius $s_{m,p}$ remains constant.[4] This element is defined as the intersection arc on the transducer surface between two spherical surfaces with radius $s_{m-1,p}$ and $s_{m,p}$, respectively, centered at point p. The computational accuracy of this single integral approach can be achieved by selecting a sufficiently narrow strip. Therefore, its computational time would depend on both the transducer radius and wavelength.

In this work, an efficient calculation method has been developed which is well suited to the determination of the acoustic pressure fields from planar circular transducers. By using the Hankel transform, the final solution can be expressed as a one-dimensional integral in which both the wavelength of the emitted ultrasound and the transducer's radius are mere parameters. To evaluate the accuracy of our integral solution, the two-dimensional surface and contour plots of acoustic pressure fields for different a/λ values have been studied and compared with those proposed by other previous investigators.[2,3] The results demonstrate that the computational time of calculating the velocity potential and its corresponding pressure field can be dramatically reduced.

I. THEORY AND METHODOLOGY

For an irrotational particle in simple harmonic motion oscillating around its rest position, caused by a vibrating planar ultrasound transducer, the particle velocity can be expressed as the gradient of a scalar velocity potential φ which satisfies the wave equation,

[a)]Author to whom correspondence should be addressed. Electronic mail: chung@lotus.mc.ntu.edu.tw

 0001-4966/99/105(4)/2243/5/$15.00 © 1999 Acoustical Society of America

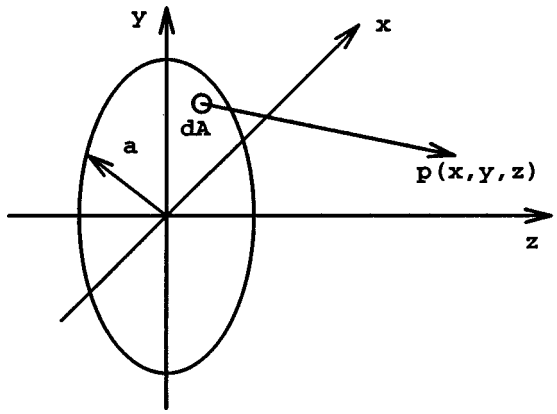


FIG. 1. Schematic diagram of a planar circular transducer used in calculating acoustic pressure fields. The transducer axial axis corresponds to the z coordinate, and the transverse axis is the r coordinate.

$$\frac{\partial^2 \varphi}{\partial t^2} = c^2 \nabla^2 \varphi, \tag{1}$$

where c represents the propagation velocity of the wave in the medium and ∇^2 is the three-dimensional Laplace operator. The solution of Eq. (1) can be written as, $\varphi = e^{j\omega t}\phi$, where ϕ is independent of time and ω is the angular frequency of the harmonic motion. Consequently, ϕ satisfies the Helmoltz equation $\nabla^2\phi + \kappa^2\phi = 0$ where $\kappa = \omega/c$ is the wave number of the propagation wave. The boundary conditions are

$$\begin{aligned} &\frac{\partial \phi}{\partial z} = f(r) \quad \text{on the } z=0 \text{ plane} \\ &\text{and} \\ &\phi(r,z) \to 0 \quad \text{as } z \to \infty. \end{aligned} \tag{2}$$

By introducing the Hankel transform, the potential function ϕ is defined by the following equations,

$$\Phi(\zeta,z) = \int_0^\infty \mathbf{r} \mathbf{J_0}(\zeta \mathbf{r}) \phi(\mathbf{r},\mathbf{z}) \mathbf{dr}$$

and

$$\phi(\mathbf{r},\mathbf{z}) = \int_0^\infty \zeta \mathbf{J_0}(\zeta \mathbf{r}) \Phi(\zeta,z) \mathbf{d}\zeta,$$

where $\mathbf{J}_0(\zeta r)$ is the zero-order Bessel function of the first kind. The Helmoltz equation becomes $\Phi_{zz} - \zeta^2\Phi + \kappa^2\Phi = 0$, under the assumption that both $\sqrt{r}\phi$ and $\sqrt{r}\phi_r$ vanish at infinity. Similarly, by applying the Hankel transform to the boundary conditions, one can obtain

$$\begin{aligned} &\phi(r,z) \to 0 \quad \text{as } z \to \infty \\ &\int_0^\infty r J_0(\zeta r) \phi \, dr = \Phi(\zeta,z) \to 0 \quad \text{as } z \to \infty, \end{aligned} \tag{3}$$

and

$$\left.\frac{\partial \phi}{\partial z}\right|_{z=0} = f(r)$$

$$\int_0^\infty \mathbf{r} \mathbf{J}_0(\zeta \mathbf{r}) \left[\left.\frac{\partial \phi}{\partial z}\right|_{z=0} \right] \mathbf{dr} = \int_0^\infty \mathbf{r} \mathbf{J}_0(\zeta \mathbf{r}) \mathbf{f}(\mathbf{r}) \mathbf{dr} = \mathbf{F}(\zeta),$$

$$\Phi_z(\zeta,0) = \mathbf{F}(\zeta).$$

The final solution of $\Phi(\zeta,z)$ can be expressed as

$$\Phi(\zeta,z) = \frac{-\mathbf{F}(\zeta)}{\sqrt{\zeta^2 - \kappa^2}} e^{-z\sqrt{\zeta^2-\kappa^2}}. \tag{4}$$

Putting the inverse Hankel transform into Eq. (4) leads to

$$\begin{aligned} \phi(\mathbf{r},z) &= \int_0^\infty \zeta \mathbf{J}_0(\zeta \mathbf{r}) \frac{-\mathbf{F}(\zeta)}{\sqrt{\zeta^2-\kappa^2}} e^{-z\sqrt{\zeta^2-\kappa^2}} \mathbf{d}\zeta \\ &= -\int_0^\infty \rho \mathbf{J}_0(\rho\zeta) \mathbf{f}(\rho) \mathbf{d}\rho \int_0^\infty \zeta \mathbf{J}_0(\zeta \mathbf{r}) \frac{e^{-z\sqrt{\zeta^2-\kappa^2}}}{\sqrt{\zeta^2-\kappa^2}} \mathbf{d}\zeta, \end{aligned} \tag{5}$$

where ρ is the distance from the origin on the transducer surface. Consider a rigid planar circular transducer with the following particle velocity,

$$\begin{aligned} &f(\rho) = U \quad \text{for} \quad 0 \leqslant \rho \leqslant a, \\ &f(\rho) = 0 \quad \text{for} \quad a < \rho. \end{aligned}$$

Equation (5) is then reduced to

$$\phi(r,z) = -\mathbf{U}\mathbf{a} \int_0^\infty \mathbf{J}_0(\zeta \mathbf{r}) \mathbf{J}_1(\zeta \mathbf{a}) \frac{e^{-z\sqrt{\zeta^2-\kappa^2}}}{\sqrt{\zeta^2-\kappa^2}} \mathbf{d}\zeta, \tag{6}$$

where $\mathbf{J}_1(\zeta a)$ is the first-order Bessel function of the first kind. Introducing a variable of integration ψ, and a variable R such that

$$R^2 = (a - re^{-j\psi})(a - re^{j\psi}), \tag{7}$$

Eq. (6) yields

$$\phi(\mathbf{r},z) = -\frac{\mathbf{U}\mathbf{a}}{2\pi} \int_{-\pi}^{\pi} \int_0^\infty \mathbf{J}_1(\zeta \mathbf{R}) \frac{e^{-z\sqrt{\zeta^2-\kappa^2}}}{\sqrt{\zeta^2-\kappa^2}} d\zeta \frac{(\mathbf{a} - \mathbf{r}\mathbf{e}^{-j\psi})}{\mathbf{R}} d\psi. \tag{8}$$

From the Hankel transform table,[5] the integration of Eq. (8) can be reduced as

$$\begin{aligned} \phi(r,z) &= \frac{Uaj}{2\pi\kappa} \int_{-\pi}^{\pi} \frac{(a - re^{-j\psi})}{R^2} (e^{-j\kappa z} - e^{-j\kappa\sqrt{R^2+z^2}}) d\psi \\ &= \frac{Uaj}{2\pi\kappa} \int_{-\pi}^{\pi} \frac{(e^{-j\kappa z} - e^{-j\kappa\sqrt{R^2+z^2}})}{a - re^{j\psi}} d\psi. \end{aligned} \tag{9}$$

Thus, the acoustic pressure p at a point (r,z) within the radiated field can be calculated by

$$p(r,z,t) = \frac{-\sigma \partial \varphi}{\partial t} = -j\sigma\omega e^{j\omega t} \phi(r,z), \tag{10}$$

where σ represents the medium's density.

II. SIMULATION RESULTS

The acoustic pressure field of a planar circular ultrasound transducer is calculated from Eqs. (9) and (10) without

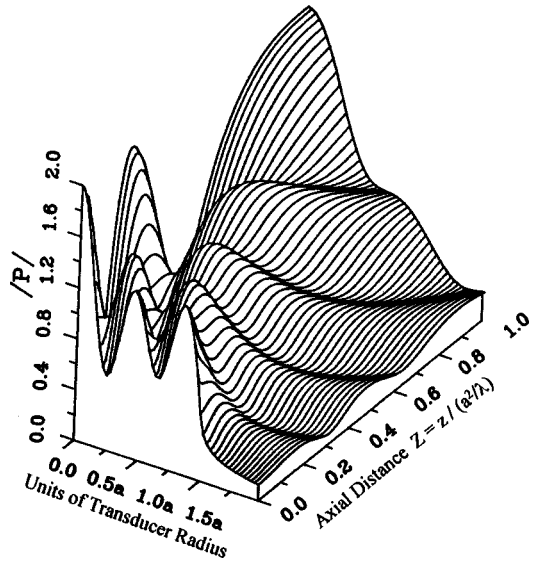


FIG. 2. Two-dimensional surface plot of acoustic pressure field for a transducer of radius 2.5 cm and frequency 0.15 MHz ($a/\lambda = 2.5$).

approximation. The acoustic pressure magnitude (p) at a point (r,z) is computed numerically on a personal computer (PC486). Owing to the symmetry, the acoustic pressure field is evaluated in the x-z plane. The sound speed (c) in the medium is 1500 m/s and the transducer's radius is 2.5 cm. Four frequencies 0.15, 0.3, 0.5, and 1.0 MHz, are selected such that the corresponding ratios of radius to wavelength in this study are 2.5, 5.0, 8.33, and 16.66, respectively.

To compare the computational speed of our integration scheme to those of the previously published methods, plots of $a/\lambda = 2.5$ and 5.0 for circular transducers are presented. A qualitative evaluation against the results of Zemanek and those of Lockwood and Willette is made. Figures 2 and 3 present the two-dimensional plots of pressure amplitude in the x-z plane for $a/\lambda = 2.5$ and 5.0, respectively. A large

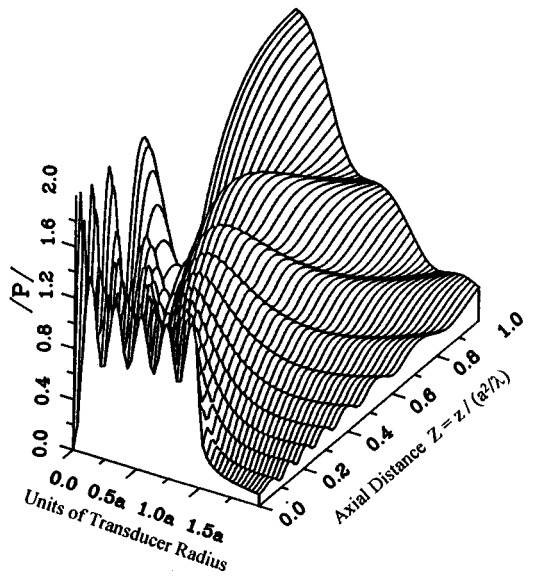


FIG. 3. Two-dimensional surface plot of acoustic pressure field for a transducer of radius 2.5 cm and frequency 0.3 MHz ($a/\lambda = 5.0$).

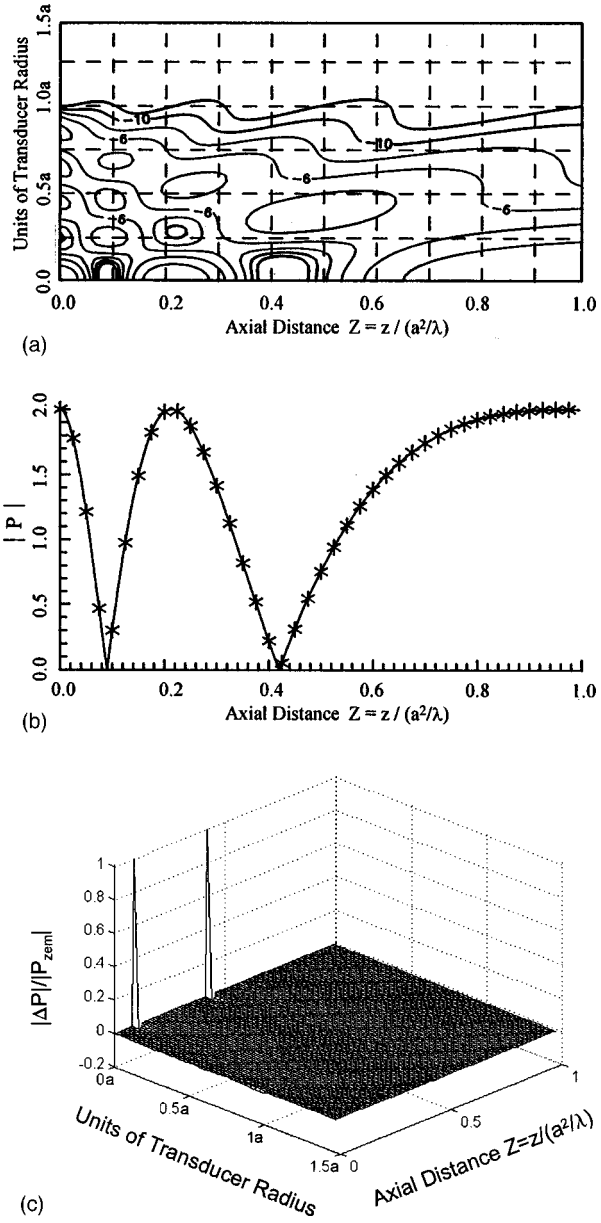


FIG. 4. (a) Acoustic pressure (dB) contour in near field of circular transducer ($a/\lambda = 2.5$). (b) Variation of pressure magnitude ($|p|$) on the transducer axis. (c) Percent difference of pressure magnitude ($|\Delta p|/|p_{zem}|$) between the acoustic pressure fields for the present one-dimensional integral scheme and Zemanek's two-dimensional integration.

number of increments (360) in ψ [Eq. (9)] is used when computing the acoustic pressure field. These two plots are compared with the plots by Zemanek [Figs. 5 and 6 of Ref. 2] and by Lockwood and Willette [Figs. 7 and 8 of Ref. 3]. Furthermore, detailed near-field pressure contours and magnitudes of the on-axis pressure variation are shown in Figs. 4(a) and (b) and 5(a) and (b) for $a/\lambda = 2.5$ and 5.0, respectively. Figures 4(c) and 5(c) are the percent difference of pressure magnitude between the acoustic pressure fields for the present one-dimensional integral scheme and the Zemanek's double integration. The comparison of the pressure contour plots and the plots for the percent difference of pressure magnitude reveals that discrepancies exist in the region next to the transducer's face and in the shape of the contour lines for small pressure values. It was partially due to the

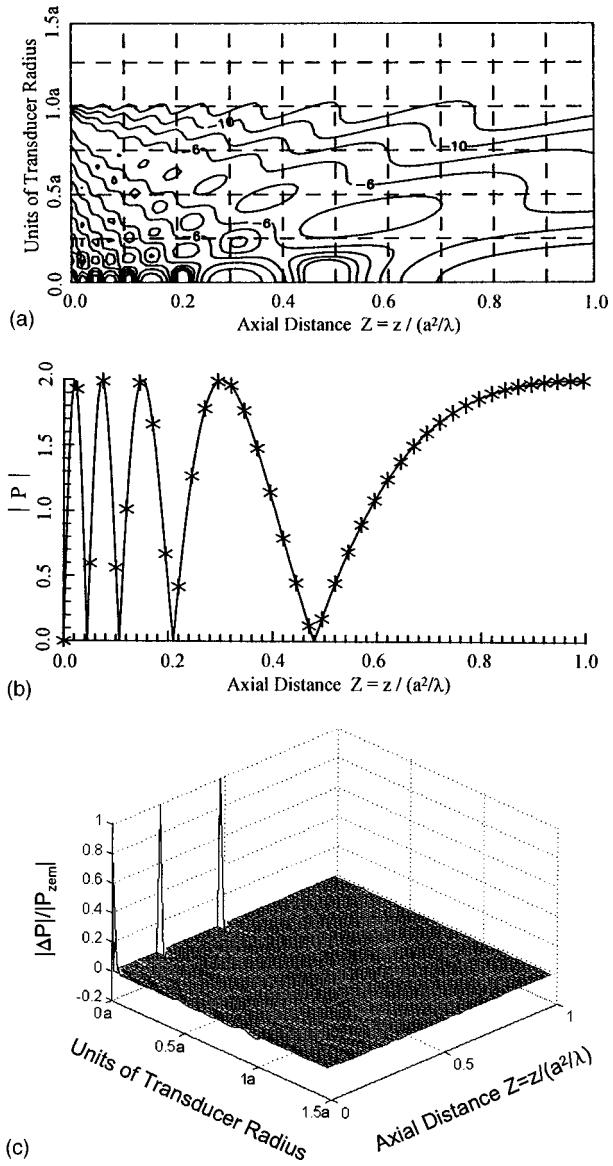


FIG. 5. (a) Acoustic pressure (dB) contour in near field of circular transducer ($a/\lambda = 5.0$). (b) Variation of pressure magnitude ($|p|$) on the transducer axis. (c) Percent difference of pressure magnitude ($|\Delta p|/|p_{\text{zem}}|$) between the acoustic pressure fields for the present one-dimensional integral scheme and Zemanek's two-dimensional integration.

numerical integration error and the way the contour lines were drawn. However, our plots and those corresponding plots of Zemanek, as well as Lockwood and Willette, appear to correlate well.

To study the effect of the number of increments in ψ [Eq. (9)] on the accuracy of the numerical scheme, the number of increments equal to 50, 100, 150, 200, 250, 300, 320, 340, and 360 are selected. The number of nodes in the x-z plane used to numerically calculate the acoustic pressure is 50×200. The total discrepancy is defined as

$$\text{DIF} = \sqrt{\frac{1}{N_i N_k} \sum_{i=1}^{N_i=50} \sum_{k=1}^{N_k=200} [p_{360}(x_i, z_k) - p_{\psi n}(x_i, z_k)]^2}, \tag{11}$$

where p_{360} and $p_{\Psi n}$ are the subsequent acoustic pressure fields for the number of increments in ψ equal to 360 and

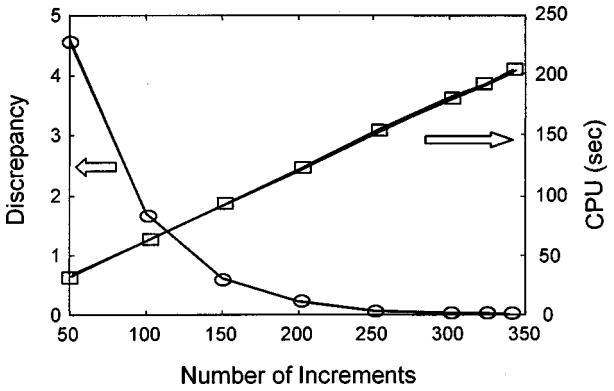


FIG. 6. Magnitude of discrepancy (curve with circle symbol) and corresponding CPU time (curve with square symbol) consumed in calculating Eq. (11) as a function of number of increments.

ψ_n, respectively. Figure 6 plots the total discrepancy and the corresponding CPU time in calculating the acoustic pressure field as a function of the number of increments for the transducer of radius 2.5 cm with frequencies 0.15, 0.3, 0.5 and 1.0 MHz. The results show that four curves overlap one another. This finding indicates that the discrepancy is independent of the excitation frequency. That is, for a given increment in ψ, the accuracy of our scheme does not vary apparently with the excitation frequency. The results also indicate that for all frequencies, our method is faster by a factor of 1.8 when using 200 increments rather than 360 without a significant loss in accuracy.

After showing the consistent results between the present scheme and the previous methods proposed by other investigators, the difference in calculating speed for the acoustic pressure fields is quantitatively evaluated. Zemanek mentioned that for a given radius and wavelength, the minimum number of elemental areas whose contributions were summed to yield a convergent result was $32\pi(a/\lambda)^2$.[2] In the present scheme, the acoustic pressure at field point p is calculated numerically from Eqs. (9) and (10). As found in Eq. (9), both the transducer radius and its wavelength are mere parameters in the integral. For a given radius and wavelength, the number of steps required for the integration depends on the number of increments in ψ. In this comparison, the number of increments in ψ is 360. The total CPU time consumed in calculating the acoustic pressure field on the x-z plane for these two schemes is further studied. Several cases with different values of a/λ (from 5.0 to 83.33) are selected: the transducer radius is 2.5 cm and the frequencies are 0.3, 0.5, 1.0, 2.0, and 5.0 MHz. Figure 7 shows the computational time consumed in calculating the acoustic pressure field for the present single-integral scheme and Zemanek's double integral. Results show that our single-integral scheme was approximately two times faster than Zemanek's for the frequency 0.3 MHz and even 500 times faster for a higher frequency, 5.0 MHz. Lockwood and Willette[3] claimed that the speed of their numerical scheme was proportional to the ratio of radius to wavelength (a/λ). Similarly, the numerical scheme presented by Roemer *et al.*[4] was also proportional to the ratio of radius to wavelength (a/λ) because the intersect strip should be sufficiently narrow for accuracy. Therefore,

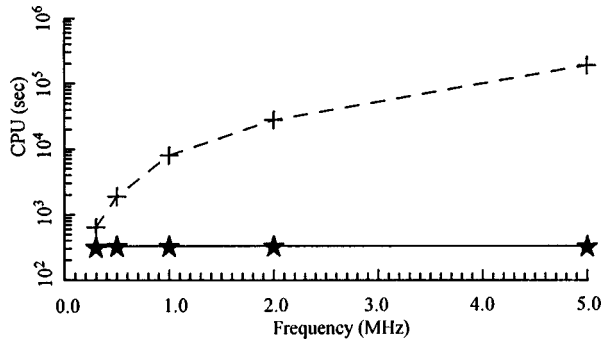


FIG. 7. CPU time consumed in calculating acoustic pressure fields of a 2.5-cm radius transducer, excited at 0.3, 0.5, 1.0, 2.0, and 5.0 MHz for two schemes; one is our proposed one-dimensional integral scheme denoted by $-*-$, and the other is Zemanek's two-dimensional integration denoted by $-+-$.

the calculating speed of our method is faster than those of the previous methods.

Figures 2–5 exhibit several maxima and zeros on the transducer axis, and gradually decay after the last maxima. The number of maxima and zeros is proportional to a/λ. Figures 2 and 3 illustrate that the pressure distribution on the r axis close to the excited surface had the same number of maxima as the pressure distribution on the transducer axis. Therefore, the pressure profile on the r axis for a high frequency (for example, 5.0 MHz) becomes a step function. The pressure magnitude in the off-axis field for low-frequency cases tends to approach zero faster at a larger distance away from the excited surface. Results show that although the numbers of maxima and zeros on axis are equal for a given value of a/λ, the corresponding maxima and zero positions in the x-z plane are quite different, and thus the resulting pressure field in the x-z plane is also different. This finding suggests that the value of a/λ cannot be used to determine the transducer size.

III. CONCLUSION

The derivation and evaluation of a single-integral scheme for computing acoustic pressure fields of planar circular ultrasound transducers have been presented. Apparently, owing to its independence of wavelength (λ) in the integral shown in Eq. (9), it has a promising advantage in quick calculation of the acoustic pressure fields for large transducers with high frequencies. The results show that our scheme has certain advantage in computing the near-field acoustic pressure distribution because of its speed and accuracy. Therefore, to calculate the pressure field for high frequency and/or large planar circular ultrasound transducers, the present single-integral scheme becomes very useful.

ACKNOWLEDGMENT

The authors would like to thank the National Science Council of the Republic of China for supporting this research under Contract No. NSC-86-2212-E168-003.

[1] L. E. Kinsler and A. R. Frey, *Fundamentals of Acoustics* (Wiley, New York, 1962).
[2] J. Zemanek, ''Beam behavior within the nearfield of a vibrating piston,'' J. Acoust. Soc. Am. **49**, 181–191 (1971).
[3] J. C. Lockwood and J. G. Willette, ''High-speed method for computing the exact solution for the pressure variations in the nearfield of a baffled piston,'' J. Acoust. Soc. Am. **53**, 735–741 (1973).
[4] R. B. Roemer, W. Swindell, S. T. Clegg, and R. Kress, ''Simulations of focused, scanned ultrasonic heating of deep-seated tumors: The effect of blood perfusion,'' IEEE Trans. Sonics Ultrason. **31**, 526–531 (1984).
[5] A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Integrals And Series: Volume 2 Special Functions* (Gordon and Breach Science, New York, 1986).

Biased lead zirconate titanate as a high-power transduction material

Mark B. Moffett[a)]
Georgia Tech Research Institute, Atlanta, Georgia 30318

Michael D. Jevnager
Newport Division, Naval Undersea Warfare Center, Newport, Rhode Island 02841

Stephen S. Gilardi
Interface Engineering, Mystic, Connecticut 06355

James M. Powers
Newport Division, Naval Undersea Warfare Center, Newport, Rhode Island 02841

(Received 19 October 1998; accepted for publication 19 January 1999)

High-power, underwater transducers utilizing polarized piezoelectric ceramic material are usually limited in drive amplitude so that depolarization does not occur, but application of a dc bias field in the polarization direction allows the use of higher ac drive fields. To demonstrate the feasibility of biased operation as a means of achieving higher power, a thin-walled, spherical-shell transducer was constructed of Channel 5800 and tested in NUWC's Acoustic Pressure Tank Facility at a hydrostatic pressure of 1400 psig (9.65 MPa). The transducer was successfully driven to 33 V/mil (1.3 MV/m) rms with an accompanying bias field of 31 V/mil (1.2 MV/m). The source level was 206 dB *re*: 1 μPa·m at 61 kHz, corresponding to a 10-dB improvement over the unbiased drive limit of 10 V/mil (0.4 MV/m) rms and 196 dB *re*: 1 μPa·m. © *1999 Acoustical Society of America.*
[S0001-4966(99)04204-6]

PACS numbers: 43.38.Ar, 43.38.Fx [SLE]

INTRODUCTION

In their quest for higher-power densities from actuators and underwater acoustic projectors, transducer designers have been led to the use of newer electrostrictive and magnetostrictive materials, such as lead magnesium niobate[1] and Terfenol-D,[2] and, more recently, single-crystal transduction materials, such as lead zinc niobate.[3] In operation, each of these new materials requires a dc bias field. Electrostrictive and magnetostrictive transducers must be biased for linearity, and the new single-crystal materials need a bias field to prevent depolarization. Older transduction materials, such as lead zirconate titanate (PZT), are not usually biased. Instead, the drive amplitude is kept within limits that have become accepted engineering practice. For example, the drive field for PZT-8 (Navy Type III) transducers has traditionally been limited to 10 V/mil (390 kV/m) rms.[4] However, in the light of the need to bias the newer materials, it seemed reasonable to consider the possibility of biasing PZT as a means of achieving higher power.[5]

The effect of a bias field can be visualized by examination of Fig. 1, which illustrates qualitatively the mechanical strain that results during and after the polarization process. Starting with the unpolarized material at the origin, the strain follows the dashed curve as the electric field intensity is increased to that required for polarization (2 MV/m, in this illustration). Further increases in electric field result in the linear behavior shown in the upper, right portion of the plot. Then, as the electric field is removed, the strain decreases along the solid curve to the remanent value, indicated by the left hand asterisk in Fig. 1. Subsequent, conventional operation of the polarized material will entail oscillations about this point, between the left and center vertical dashes, which are considered to be practical drive limits for unbiased operation. (In Fig. 1, we have shown these limits as ± 0.5 MV/m as an illustration.) Note that the lower limit, indicated by the left vertical dash, is near the knee of the depolarization curve. Larger amplitudes would involve partial depolarization of the material and increasing losses due to an enlarging hysteresis loop. Therefore, one avoids fields having negative excursions to the left of the left vertical dash. On the other hand, positive electric drive fields greater than that indicated by the center vertical dash would not entail a reduction of the polarization state, but would, instead, help to maintain it. For example, operation between the extreme left and right vertical dash points of Fig. 1, around the bias point indicated by the right-hand asterisk (1 MV/m in this illustration), entails an amplitude three times larger than would be possible without the dc bias field.

Baerwald and Berlincourt[6] discussed the advantages of biasing barium titanate piezoelectric ceramic material. For continuous operation, in which the heating due to losses associated with a hysteresis loop can cause the temperature to exceed the Curie point, they showed that biasing could reduce the hysteresis loop area, achieving temperature stabilization, and allowing higher drive fields. Hueter *et al.*[7] followed up on the loss-factor reduction aspects of Ref. 6, but

[a)]Currently on Intergovernmental Personnel Act (IPA) assignment at NUWC, Newport, RI 02841.

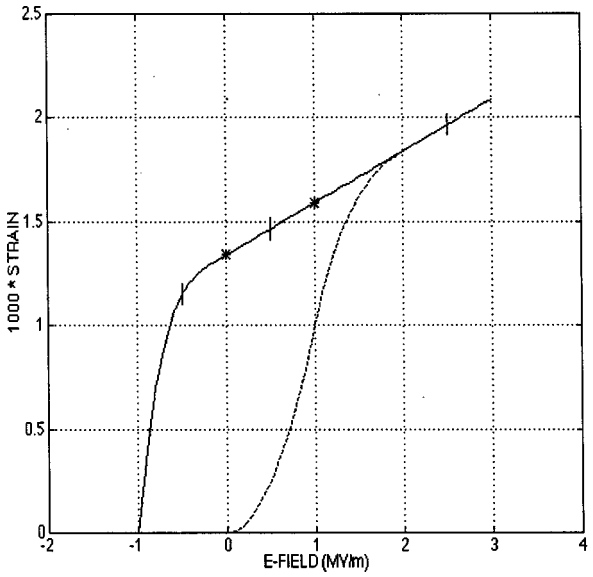


FIG. 1. Hypothetical plot of strain vs E-field for piezoelectric ceramic material. The lower (dashed) curve depicts the initial application of the polarization voltage. The upper curve is the result for the polarized (i.e., piezoelectric) material. The left pair of vertical dashes corresponds to the driving field limits for conventional, unbiased operation around the zero-field point (indicated by the left-hand asterisk). For operation about the bias point (denoted by the right-hand asterisk), the driving field oscillates between the limits indicated by the extreme left and right vertical dashes.

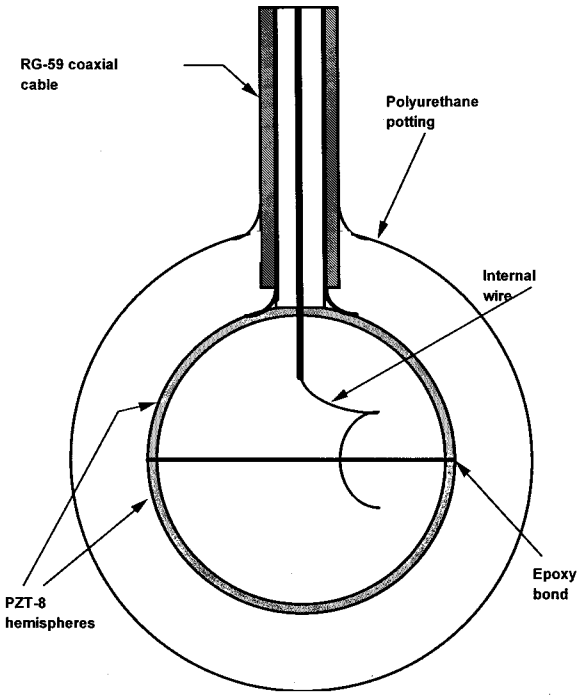


FIG. 2. Spherical-shell transducer, with coaxial cable dielectric bonded to sphere. The sphere outer diameter is 1 in., and the wall thickness is 0.037 in.

noted that when mechanical losses are dominant (in comparison with dielectric losses), biasing does not provide much benefit. For this reason, along with the advent of PZT ceramics, with their clear superiority to barium titanate, and because biased operation requires a more complicated drive system and increased dielectric insulation, the motivation for the use of biasing as a means of obtaining increased power has been lacking. Over the years since the studies of Refs. 6 and 7, however, underwater acoustic transducers have often turned out to be electric-field-limited, rather than mechanical-stress-limited.[4,8] Under these conditions, biasing can be expected to allow the use of higher drive fields and concomitant higher acoustic energy densities. However, some transducers, such as high-Q projectors, whose maximum amplitude near resonance is dictated by mechanical stress limits, will not benefit from biased operation.

I. EXPERIMENT

To demonstrate that biased operation of a PZT transducer will allow larger sound pressures to be produced than are possible with conventional, unbiased operation, the thin-walled, spherical-shell transducer illustrated schematically in Fig. 2 was constructed from two hemispheres of Channel 5800 piezoelectric ceramic material, a hard PZT with properties similar to those of Navy Type III.[9] The outer electrode was removed near the edge of each hemisphere to form an electrode-free margin of 1.5 mm at the equator of the assembled sphere. A hole of diameter 0.8 mm was drilled in one of the hemispheres for insertion of the RG-59 coaxial cable's solid center conductor. After removal of the outer electrode from a 10-mm-diameter circle around the hole and plasma etching[10] the end of the cable, the cable's polyethylene dielectric was epoxy-bonded to the outside surface of the hemisphere. A stranded wire was soldered at several points inside the two hemispheres before they were epoxy-bonded to form a sphere with outer diameter of 1 in. (25.4 mm) and average wall thickness of 0.037 in. (0.94 mm). After the cable shield was soldered to the outer electrode of the sphere, the assembly was encapsulated in a polyurethane (PR 1538[11]) boot. The regions that were expected to sustain large electric field intensities, near the center conductor, were filled with solid, bubble-free, dielectric insulators in order to avoid voltage breakdown, which is much more likely to occur when air is present.

Low-level acoustic measurements were performed at NUWC's Dodge Pond Facility, but the high-level testing was

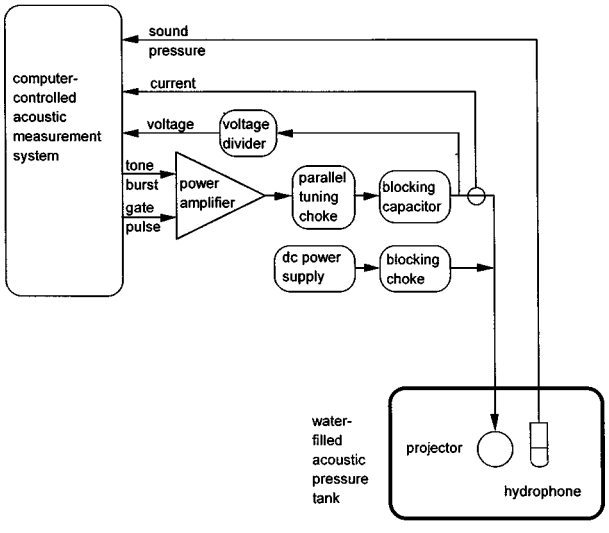


FIG. 3. Schematic of experimental setup.

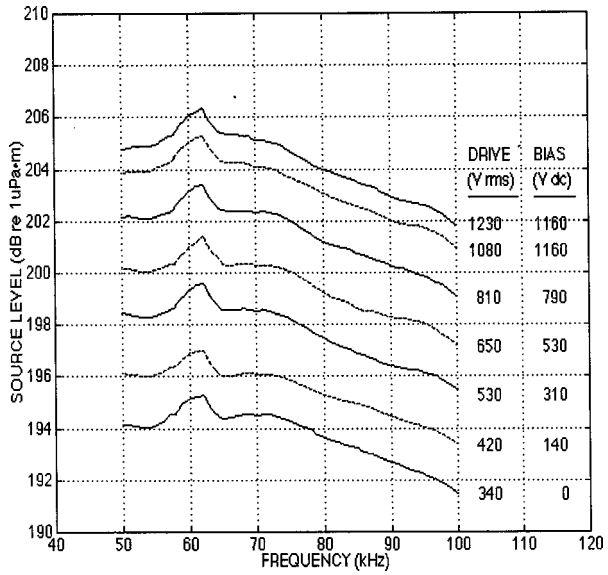


FIG. 4. Source level vs frequency for various drive and bias levels.

done in NUWC's Acoustic Pressure Tank Facility (APTF). To prevent dynamic tension in the PZT shell, compressive prestress up to 10 kpsi (70 MPa) was applied to the sphere by pressurization to 1400 psi (10 MPa). Transmitting voltage response (TVR) measurements were performed at a low ($\leqslant$0.3%) duty cycle in order to avoid undue heating of the transducer. The experimental arrangement is shown schematically in Fig. 3. Thirty cycle tone bursts, along with accompanying gate pulses, were supplied by the APTF measurement system to a 40-vacuum-tube, 20-kW, gated power amplifier.[12] The frequency was varied in 1-kHz steps from 50 to 100 kHz. For each frequency sweep, the drive voltage was allowed to vary ($\pm$1 dB) according to amplifier loading, but it was gradually increased (typically in 2-dB steps) for successive frequency sweeps. A 400-μH, parallel-tuning inductor was placed across the amplifier output, and that was followed by a series, 1-μF, blocking capacitor to prevent direct current from flowing into the tuning inductor and the amplifier output transformer. To limit the alternating voltage across the dc power supply, a 12-mH, blocking inductor was placed in series with the power supply.

II. RESULTS AND DISCUSSION

The source levels are plotted in Fig. 4, as a function of frequency, for bias voltages ranging from 0 to 1160 V. The corresponding rms drive voltages (at 50 kHz) ranged from 340 V to 1230 V, respectively. (The largest voltage occurring during a frequency sweep was always at 50 kHz, the lowest frequency.) The spherical-shell wall thickness was 37 mils (0.94 mm), and so the maximum rms drive field at 50 kHz was 33 V/mil (1.3 MV/m) with an accompanying bias field of 31 V/mil (1.2 MV/m). This drive level is 10.4 dB greater than the permitted value of 10 V/mil (0.4 MV/m) for unbiased operation. The peak (dc+1.4$\times$ac) field was 78 V/mil (3.1 MV/m).

During an attempt to exceed the 33-V/mil drive level, there occurred a voltage breakdown that apparently produced a dc, resistive shunt across the transducer. (Subsequent measurements with a megohmmeter indicated that the dc resistance was a function of the dc voltage used to measure it, decreasing from more than 10 GΩ at 10 V to less than 500 kΩ at 200 V.) This is thought to be the result of a carbon track deposited at the transducer end of the coaxial cable. At the time of this writing, the transducer had not yet been dissected to verify the cause of the shunt. In spite of this resistive shunt, the transducer can still be driven without bias at 10 V/mil, where it produces 196 dB *re*: 1 μPa·m at 61 kHz.

Table I lists the transducer performance parameters in their chronological order, from left to right. The (constant-electric-field) resonance frequency, f_r^E, is defined here as the frequency of maximum conductance, $G(f_r^E)$, and the susceptance at that frequency, $B(f_r^E)$, is equal to the blocked value. It can be seen from Table I that there was no significant change in any of these quantities as a result of the application of the bias voltage, E_0, nor did any appreciable changes result from the voltage breakdown. From the lack of bias-field dependence, we can conclude that biasing did not affect the PZT piezoelectric properties. Voltage breakdown would not be expected to affect the ac operation of the transducer as long as the dc resistance remained much larger than the ac impedance of the transducer. (However, repeated applications of dc voltage, such as use of the megohmmeter, would

TABLE I. Transducer parameters as measured before and after voltage breakdown. P_0=hydrostatic pressure, E_0=bias voltage, E_{50}=drive voltage at 50 kHz, SL_{50}=source level at 50 kHz, TVR_{50}=transmitting voltage response at 50 kHz, f_r^E=resonance frequency for constant electric field, $G(f_r^E)$=maximum value of conductance, and $B(f_r^E)$=blocked susceptance.

	Prior to voltage breakdown							After breakdown
P_0 [psig]	500	500	1000	1000	1000	1200	1400	1400
E_0 [V dc]	0	140	310	530	790	1160	1160	0
E_{50} [V rms]	340	420	530	650	810	1080	1230	360
SL_{50} [dB *re*: 1 μPa·m]	194.1	196.1	198.4	200.2	202.2	203.8	204.7	193.9
TVR_{50} [dB *re*: 1 μPa·m/V]	143.4	143.5	144.0	144.0	144.0	143.2	142.9	142.9
f_r^E [kHz]	59	59	60	59	60	61	61	63
$G(f_r^E)$ [mS]	2.7	2.8	3.0	3.0	3.0	2.7	2.5	2.7
$B(f_r^E)$ [mS]	5.8	5.8	5.6	5.6	5.6	5.5	5.5	5.2

probably cause further breakdown and carbon tracking, ultimately lowering the dc shunt resistance to the point where the ac performance would be adversely affected.)

To verify these results, a second spherical shell transducer, also made from a pair of Channel 5800 hemispheres, was constructed and tested. Unlike the first transducer, depicted in Fig. 2, in which the polyethylene cable dielectric was epoxy-bonded to the sphere, the second design involved a separation between the end of the cable dielectric and the sphere, with the intervening space filled with polyurethane (Uralite 3140) potting material. This transducer was driven, without incident and with linear behavior, to 30 V/mil (1.2 MV/m). The bias field was 33 V/mil (1.3 MV/m). Breakdown occurred at an rms drive field of 35 V/mil (1.4 MV/m), corresponding to a peak field of 82 V/mil (3.2 MV/m), and resulting in a dc, resistive shunt, similar to that described above. Again, the ac performance at 10 V/mil rms and zero bias was unaffected by the shunt.

It should be noted that (prior to breakdown) the power required for biasing (<1 mW) was negligible compared to the radiated power (>2 kW). Therefore the transducer efficiency was unaffected by biasing.

III. CONCLUSIONS AND RECOMMENDATIONS

This experiment has demonstrated the use of biased PZT as a high-power transduction material. Only one material, Channel 5800, has been tested so far, but other PZT formulations should be worthy of investigation. The results suggest that an rms drive field of 30 V/mil (1.2 MV/m) is feasible if it is accompanied by a dc bias field (in the polarization direction) sufficient to prevent negative peak fields exceeding 14 V/mil (0.6 MV/m). This will result in a positive peak field of 71 V/mil (2.8 MV/m), i.e., five times the value encountered during conventional, unbiased operation at 10 V/mil (0.4 MV/m). Thus to get the 9.5-dB benefit of a threefold increase in rms drive voltage, one must provide for sufficient insulation in the design of the transducer to prevent voltage breakdown due to a fivefold increase in the peak field. Transducer designers might want to consider PZT as an alternative to some of the newer, more exotic, but not yet fully developed, ceramic and single-crystal materials, because these materials too will have to be biased for high-power applications. Perhaps PZT, which has served so well as the material of choice in the past, will also be the material of choice in the future.

ACKNOWLEDGMENTS

The authors wish to acknowledge the contributions of Patrick J. Monahan and William Duchemin for assistance with transducer construction; Joseph M. Pascucci and Stephen G. Savitsky for low-level transducer measurements at Dodge Pond; William L. Clay and William L. Konrad (Omni Engineering) for guidance in the operation of the power amplifier; John M. Armstrong, Christopher A. Wyatt, and Robert M. Drake for assistance with the APTF test setup; and Anthony E. Paolero and Roger R. St. Martin (General Physics Federal Systems) for conducting the high-level measurements. (Except for Konrad and St. Martin, all were employees of NUWC.) This work was sponsored by the Office of Naval Research, Code 321.

[1] K. Uchino, ''Electrostrictive actuators: Materials and applications,'' Ceram. Bull. **65**, 647–652 (1986).
[2] M. B. Moffett, A. E. Clark, M. Wun-Fogle, J. F. Lindberg, J. P. Teter, and E. A. McLaughlin, ''Characterization of Terfenol-D for magnetostrictive transducers,'' J. Acoust. Soc. Am. **89**, 1448–1455 (1991).
[3] S.-E. Park and T. R. Shrout, ''Characteristics of relaxor-based piezoelectric single crystals for ultrasonic transducers,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **44**, 1140–1147 (1997).
[4] R. S. Woollett, ''Power limitations of sonic transducers,'' IEEE Trans. Sonics Ultrason. **SU-15**, 218–229 (1968).
[5] M. B. Moffett, J. M. Powers, R. S. Janus, and E. A. McLaughlin, ''Biased lead zirconate-titanate as a high-power transducer material,'' NUWC-NPT Reprint Report 10,766 (15 May 1997), presented at 1997 ONR Transducer Materials and Transducers Workshop, 29 April–1 May 1997, Pennsylvania State University, University Park, PA.
[6] H. G. Baerwald and D. A. Berlincourt, ''Electromechanical response and dielectric loss of prepolarized barium titanate under maintained electric bias. Part I,'' J. Acoust. Soc. Am. **25**, 703–710 (1953).
[7] T. F. Hueter, D. P. Neuhaus, and J. Kolb, ''An experimental study of polarization effects in barium titanate ceramics,'' J. Acoust. Soc. Am. **26**, 696–703 (1954).
[8] M. B. Moffett, ''On the power limitations of sonic transducers,'' J. Acoust. Soc. Am. **94**, 3503–3505 (1993).
[9] Channel Industries, Inc., ''Piezoelectric Ceramics'' (839 Ward Drive, Santa Barbara, CA 93111), pp. 14–15.
[10] Plasma-etching done by DDU Enterprises, Suite A2, 2909 Oregon Ct., Torrance, CA 90503.
[11] Product of Courtaulds Aerospace, Inc., 5430 San Fernando Rd., Glendale, CA 91203.
[12] Power amplifier designed by William L. Konrad.

Sensor coupling in acoustic media using reciprocity

Johan Vos, Guy G. Drijkoningen, and Jacob T. Fokkema
Delft University of Technology, 2600 GA Delft, The Netherlands

(Received 20 July 1998; accepted for publication 31 December 1998)

In this paper the coupling of a sensor to an acoustic medium is discussed. Based on the acoustic reciprocity theorem, an expression is derived for the motion of the sensor as a function of the undisturbed motion of the embedding medium. What is special here is that the sensor *is* the scattering object. The sensor coupling is affected by two factors: the ratio of the density of the sensor and of the embedding medium (ground), and a frequency-dependent factor depending on the geometry of the sensor. © *1999 Acoustical Society of America.* [S0001-4966(99)02204-3]

PACS numbers: 43.38.Ar, 43.20.Tb [SLE]

INTRODUCTION

A sensor is a device that converts a physical quantity to an electric signal. In this article focus is on motion-sensitive and pressure-sensitive sensors. What the sensor is meant to measure is the motion of the medium if the sensor were not there at all. This can never be the case since the measuring device itself is present. When an acoustic wave arrives at a contrasting medium—e.g., a measuring device—a reflected wave will exist. As a consequence, the original acoustic wavefield will be disturbed. Since it is the goal of measuring devices to obtain values for the original acoustic wave field—the wave field that would exist if no device was present—it is important to have an idea about the magnitude and behavior of the disturbing acoustic field. The fact that a device meant for the detection of an acoustic wave field disturbs this wave field exists in general. In this article, focus is given on the specific case of an acoustic device—a pressure or velocity measuring sensor, embedded *in* a medium—e.g., the earth. It must be stressed here that although the examples in this paper are taken from seismological practice, the theoretical results have an impact on many more fields of application, like in nondestructive testing, underwater acoustics, etc. It will be shown here that if true amplitudes for any acoustical application are desired, a correction must be made for the coupling of the sensor to the probing medium, the amount of which is quantified in this paper.

The sensor coupling is defined to be the ratio between the velocity of the sensor (v_{sens}), and the velocity of the medium (v_{med}) in absence of the sensor:

$$C_{\text{SC}}=\frac{v_{\text{sens}}}{v_{\text{med}}}. \tag{1}$$

In the past, different models were used in order to describe the problem of the sensor coupling. This research was started to address the problem of the sensor coupling as used in seismic exploration, where the sensor on land is the geophone. The geophone is an electro-magnetic device, which is put on top of a spike which furnishes the coupling to the ground. In most theoretical studies for geophones, the sensor was simulated by its case only, neglecting the spike. Tan (1987) was the first to notice that the spike is too important to ignore. In this study, the influence of the spike will be examined. In practice, spikes of geophones are made of steel, so in the analysis the sensor is considered to be a rigid, movable object. It must be stressed here that although these problems are taken from a geophysical practice, the results here pertain to a much more general audience of people who concern themselves with coupling problems.

Sensor coupling involves two phenomena: scattering due to the sensor being there, and boundary effects of the sensor due to bad contact between the sensor and the embedding medium. The first problem is called the interaction coupling, denoted by C_I, and the second contact coupling, denoted by C_C. Then the total sensor coupling can be written as:

$$C_{\text{SC}}=\frac{v_{\text{sens},pc}}{v_{\text{med}}}\,\frac{v_{\text{sens}}}{v_{\text{sens},pc}}=C_I C_C. \tag{2}$$

In the case of interaction coupling, no slip or friction can occur, and the particle velocity shows a continuous behavior when approaching the sensor. Because of differences in material properties between the medium and the sensor, material displacement in absence of the sensor is not exactly the same as the sensor displacement. The sensor can be considered to be a contrasting domain in a half-space, e.g., the earth. In practice, however, slip and friction may occur, and the sensor will not exactly follow the motion of the material in the immediate vicinity of the sensor. This effect is separated in the above by the factor C_C. As an extreme case of this situation, the sensor may be in completely loose contact with the medium. In this paper this situation will not be addressed; only the interaction coupling will be considered, where perfect contact between the sensor and the embedding medium is assumed. In this case the disturbance of the wave field, due to the presence of the sensor, has to be investigated.

The paper is built up as follows. First, a homogeneous half-space is considered, and then the effect of a contrasting domain in the half-space is formulated. The special case of a cylindrical sensor will be considered for its analytical convenience: Most of the integrals can be determined analytically. Also, a block-shaped contrast will be analyzed where some integrals can also be determined analytically. A few numerical results will be shown.

I. THEORY

A. Basic acoustic equations

An acoustic medium is considered with constitutive coefficients ρ and κ, where ρ is the volume density of mass, given in $\mathrm{kg\,m^{-3}}$, and κ is the compressibility, given in $\mathrm{Pa^{-1}}$. The wave field in a point of an acoustic medium is fully defined by its pressure p and the three components of the particle velocity v_k. In the medium, two kinds of sources may be present: monopoles and dipoles. A monopole source is described by q, the volume source density of injection rate [1/s]. A dipole source is quantified by f_k, the volume source density of volume force ($\mathrm{N\,m^{-3}}$). In the notation of these quantities and the wave field description, the notation of Fokkema and Van den Berg (1993) is closely followed.

With these quantities, the basic equations describing wave motion are given by Newton's second law, the equation of motion,

$$\partial_k p + \rho\, \partial_t v_k = f_k, \tag{3}$$

and by the deformation equation, using a linear, instantaneously reacting isotropic medium in the low-velocity approximation:

$$\partial_k v_k + \kappa\, \partial_t p = q, \tag{4}$$

where ∂_t stands for differentiation with respect to time, and ∂_k denotes differentiation with respect to the spatial coordinate x_k. The analysis is performed in the Laplace domain, where the Laplace transformation is given by

$$\hat{\phi}(s) = \int_0^\infty \exp(-st)\,\phi(t)\,dt, \tag{5}$$

where ϕ can be any quantity, a hat on top of a quantity denotes that it is in its Laplace-domain representation, and s is the Laplace transform parameter. By applying the one-sided Laplace transformation with respect to time, the above equations are written as

$$\partial_k \hat{p} + s\rho \hat{v}_k = \hat{f}_k \tag{6}$$

and

$$\partial_k \hat{v}_k + s\kappa \hat{p} = \hat{q}, \tag{7}$$

where temporal initial conditions of the wave field are taken into account in the sources.

B. Homogeneous half-space

A homogeneous acoustic half-space is considered. The half-space is of infinite extent in both sides for the x_1 and x_2 directions. The x_3 axis is perpendicular to the bounding surface, and points downward into the half-space. Since there is a contrasting plane, the incident waves will be reflected at the boundary of the half-space, giving rise to a scattered wave field. From Fokkema and Van den Berg (1993), it follows that the total wave field due to a monopole source in a homogeneous acoustic half-space is given by

$$\hat{p}^H(\mathbf{x},s) = \hat{q}^S(s)\hat{G}^H(\mathbf{x}^S|\mathbf{x},s), \tag{8}$$

and

$$\hat{v}_k^H(\mathbf{x},s) = -\hat{q}^S(s)\hat{\Gamma}_k^H(\mathbf{x}^S|\mathbf{x},s). \tag{9}$$

In these expression the Green's functions are introduced for a homogeneous half-space, $\hat{G}^H$ and $\hat{\Gamma}_k^H$, for the pressure and the particle velocity, respectively; the superscript H refers to the half-space, and the superscript S refers to the source position. The Green's functions are given by

$$\hat{G}^H(\mathbf{x}^S|\mathbf{x},s) = \hat{G}(\mathbf{x}-\mathbf{x}^S,s) - \hat{G}(\mathbf{x}-\mathbf{x}^{S*},s) \tag{10}$$

and

$$\hat{\Gamma}_k^H(\mathbf{x}^S|\mathbf{x},s) = \partial_k(\hat{G}(\mathbf{x}-\mathbf{x}^S,s) - \hat{G}(\mathbf{x}-\mathbf{x}^{S*},s)) \tag{11}$$

with $\hat{G}(\mathbf{x},s)$ being defined as:

$$\hat{G}(\mathbf{x},s) = \frac{\exp\left(-\dfrac{s}{c}|\mathbf{x}|\right)}{4\pi|\mathbf{x}|}. \tag{12}$$

S^* is the mirror image of the source position S.

C. Contrasting domain in half-space

Now an acoustic half-space with a contrasting domain, e.g., a geophone-spike, is considered. The total wave field in an acoustic half-space due to a dipole source located somewhere in the medium, but outside a contrasting domain $\partial D_{\mathrm{sct}}$ can be decomposed into two wave fields: an incident wave field and a scattered wave field. The incident field is now considered to be the half-space response. Then, the total pressure is expressed by

$$\hat{p}^{\mathrm{tot}}(\mathbf{x},s) = \hat{p}^H(\mathbf{x},s) + \hat{p}^{\mathrm{sct}}(\mathbf{x},s). \tag{13}$$

In the derivations so far, the expression followed straightforwardly. In this section use is made of the acoustic reciprocity theorem as elaborated in Fokkema and Van den Berg (1993). The acoustic reciprocity theorem is a very elegant way to formulate a scattering problem, especially in geometries of laterally varying structures. In the reciprocity theorem two states are defined for the acoustic source, the material parameters and wave field quantities. In the case of a contrasting domain, the states for the sources are the same, while the states of the material parameters and the wave field quantities are different in the area of the contrasting domain.

Since the incident field is source-free in the contrasting domain, one can write for a point $\mathbf{x}^R$ outside the contrasting domain [Fokkema and Van den Berg, 1993, Eq. (8.46)]:

$$0 = \int_{x\in\partial D_{\mathrm{sct}}} [\hat{G}^H(\mathbf{x}^R|\mathbf{x},s)\hat{v}_k^H(\mathbf{x},s) + \hat{\Gamma}_k^H(\mathbf{x}^R|\mathbf{x},s)\hat{p}^H(\mathbf{x},s)]n_k\,dA, \quad \mathbf{x}^R \in D'_{\mathrm{sct}}. \tag{14}$$

The scattered field is source-free in the domain exterior to the contrasting domain, leading to [Fokkema and Van den Berg, 1993, Eq. (8.47)]

$$\hat{p}^{\mathrm{sct}}(\mathbf{x}^R,s) = \int_{x\in\partial D_{\mathrm{sct}}} [\hat{G}^H(\mathbf{x}^R|\mathbf{x},s)\hat{v}_k^{\mathrm{sct}}(\mathbf{x},s) + \hat{\Gamma}_k^H(\mathbf{x}^R|\mathbf{x},s)\hat{p}^{\mathrm{sct}}(\mathbf{x},s)]n_k\,dA, \quad \mathbf{x}^R \in D'_{\mathrm{sct}}. \tag{15}$$

The scattered field is obtained in terms of the total field by adding Eqs. (14) and (15). However, the goal is to find representations for the limiting behavior where the considered point $\mathbf{x}^R$ approaches the boundary $\partial D_{\mathrm{sct}}$. From Fokkema and Van den Berg's (1993) Eq. (8.56), it follows that the acoustic pressure in a point located at the boundary of the sensor is given by

$$\hat{p}^{\mathrm{sct}}_{\partial D_{\mathrm{sct}}}(\mathbf{x}^R,s)=\tfrac{1}{2}\hat{p}^{\mathrm{tot}}(\mathbf{x}^R,s)+\int_{\mathbf{x}\in\partial D_{\mathrm{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)}[\hat{G}^H(\mathbf{x}^R|\mathbf{x},s)\hat{v}^{\mathrm{tot}}_k(\mathbf{x},s)+\hat{\Gamma}^H_k(\mathbf{x}^R|\mathbf{x},s)\hat{p}^{\mathrm{tot}}(\mathbf{x},s)]n_k\,dA, \tag{16}$$

where integration over the scattering domain now excludes the singular point of observation, which is represented by the term $(1/2)\hat{p}^{\mathrm{tot}}$.

D. Rigid movable object

The condition that the contrast is a movable rigid object will be used now, so the particle velocity is constant in space. Therefore, this particle velocity is denoted with a capital: $\hat{V}^{\mathrm{tot}}_k$. The equation of motion applied to the scattering domain gives

$$\int_{\mathbf{x}\in\partial D}\hat{p}^{\mathrm{tot}}n_k\,dA=-s\hat{V}^{\mathrm{tot}}_k\int_{\mathbf{x}\in D}\rho_{\mathrm{sens}}\,dV \tag{17}$$

in which ρ_{sens} is the density of the sensor. An expression for the pressure $\hat{p}$ as a function of the spatial constant velocity $\hat{V}_k$ is given by

$$\hat{p}^{\mathrm{tot}}=-s\rho_{\mathrm{sens}}x_l\hat{V}^{\mathrm{tot}}_l, \tag{18}$$

which can easily be shown by substituting this solution in Eq. (17) and applying Gauss' divergence theorem. This leads to

$$\hat{V}_l\int_{\mathbf{x}\in D}\partial_k x_l\,dV=\hat{V}_k\int_{\mathbf{x}\in D}dV \tag{19}$$

and the posed solution obeys the equation of motion in global form. So then, for a rigid movable contrast, the scattered field becomes:

$$\hat{p}^{\mathrm{sct}}_{\partial D_{\mathrm{sct}}}(\mathbf{x}^R,s)=\tfrac{1}{2}\hat{p}^{\mathrm{tot}}(\mathbf{x}^R,s)+\hat{V}^{\mathrm{tot}}_k\int_{\mathbf{x}\in\partial D_{\mathrm{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)}\hat{G}^H(\mathbf{x}^R|\mathbf{x},s)n_k\,dA-s\rho_{\mathrm{sens}}\hat{V}^{\mathrm{tot}}_k\int_{\mathbf{x}\in\partial D_{\mathrm{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)}x_k\hat{\Gamma}^H_l(\mathbf{x}^R|\mathbf{x},s)\times n_l\,dA,\quad \mathbf{x}^R\in\partial D'_{\mathrm{sct}}, \tag{20}$$

where now $\hat{V}^{\mathrm{tot}}_k$ is taken outside the integral, since it is a constant along the scattering surface.

E. Coupling coefficients for one-dimensional motion

The goal of this paper is to obtain a coupling coefficient as defined in the introduction, either by $\hat{p}^{\mathrm{tot}}/\hat{p}^H$ or by $\hat{v}^{\mathrm{tot}}_k/\hat{v}^H_k$, where it is understood that the undisturbed state is the field of the half-space, denoted by the superscript H. In order to define the coupling coefficient, $\hat{p}$ needs to be converted to $\hat{v}_k$, or vice versa.

First, the ratio $\hat{p}^{\mathrm{tot}}/\hat{p}^H$ is considered. For this ratio the coefficients v^{tot}_k cannot be determined only in terms of $\hat{p}^{\mathrm{tot}}$, unless a motion only in *one* direction is assumed. Assuming only a motion in the *vertical* direction one can write

$$\hat{v}^{\mathrm{tot}}_3=-\frac{1}{s\rho_{\mathrm{sens}}x_3}\hat{p}^{\mathrm{tot}} \tag{21}$$

as follows from Eq. (18). The scattered field [Eq. (20)] is written as

$$\begin{aligned}\hat{p}^{\mathrm{tot}}(\mathbf{x}^R,s)-\hat{p}^H_{\partial D_{\mathrm{sct}}}(\mathbf{x}^R,s)&=\frac{1}{2}\hat{p}^{\mathrm{tot}}(\mathbf{x}^R,s)-\frac{1}{s\rho_{\mathrm{sens}}x^R_3}\hat{p}^{\mathrm{tot}}(\mathbf{x}^R,s)\\&\quad\times\int_{\mathbf{x}\in\partial D_{\mathrm{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)}\hat{G}^H(\mathbf{x}^R|\mathbf{x},s)n_3\,dA\\&\quad+\frac{1}{x^R_3}\hat{p}^{\mathrm{tot}}(\mathbf{x}^R,s)\int_{\mathbf{x}\in\partial D_{\mathrm{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)}x_3\hat{\Gamma}^H_l(\mathbf{x}^R|\mathbf{x},s)n_l\,dA\\&=B\hat{p}^{\mathrm{tot}}(\mathbf{x}^R,s),\quad \mathbf{x}^R\in D'_{\mathrm{sct}},\end{aligned} \tag{22}$$

where B is the factor which can be determined for any configuration:

$$\begin{aligned}B&=\frac{1}{2}-\frac{1}{s\rho_{\mathrm{sens}}x^R_3}\int_{\mathbf{x}\in\partial D_{\mathrm{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)}\hat{G}^H(\mathbf{x}^R|\mathbf{x},s)n_3\,dA\\&\quad+\frac{1}{x^R_3}\int_{\mathbf{x}\in\partial D_{\mathrm{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)}x_3\hat{\Gamma}^H_l(\mathbf{x}^R|\mathbf{x},s)n_l\,dA.\end{aligned} \tag{23}$$

So the interaction coupling coefficient C_I for pressure-sensitive sensors can then be determined as

$$C_I=\frac{\hat{p}^{\mathrm{tot}}}{\hat{p}^H}=\frac{1}{1-B}. \tag{24}$$

This analysis is valid for pressure-sensitive sensors. For sensors sensitive to particle velocity, the equations of motion are used to introduce $\hat{v}^{\mathrm{tot}}_k$, and the scattered field [Eq. (20)] can be written as

$$\begin{aligned}-s\rho_{\mathrm{sens}}x^R_k\hat{v}^{\mathrm{tot}}_k(\mathbf{x}^R,s)+s\rho x^R_k\hat{v}^H_k(\mathbf{x}^R,s)&=-\tfrac{1}{2}s\rho_{\mathrm{sens}}x^R_k\hat{v}^{\mathrm{tot}}_k(\mathbf{x}^R,s)+\hat{v}^{\mathrm{tot}}_k(\mathbf{x}^R,s)\\&\quad\times\int_{\mathbf{x}\in\partial D_{\mathrm{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)}\hat{G}^H(\mathbf{x}^R|\mathbf{x},s)n_l\,dA-s\rho_{\mathrm{sens}}\hat{v}^{\mathrm{tot}}_k(\mathbf{x}^R,s)\\&\quad\times\int_{\mathbf{x}\in\partial D_{\mathrm{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)}x_k\hat{\Gamma}^H_l(\mathbf{x}^R|\mathbf{x},s)n_l\,dA.\end{aligned} \tag{25}$$

Again, the coupling coefficient for each component $\hat{v}_k$ cannot be determined, unless one-dimensional motion is assumed. Again, say that only a motion in the vertical direction exists. Then it holds that

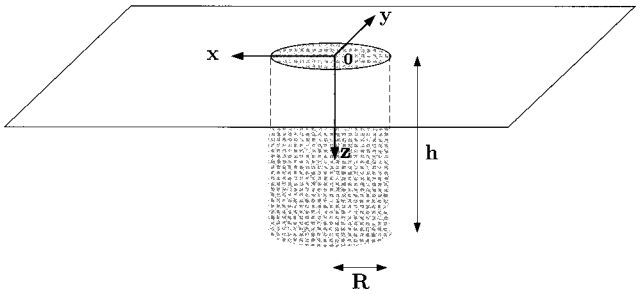


FIG. 1. Configuration for cylinder-shaped contrast.

$$
\begin{aligned}
&-s\rho_{\text{sens}}x_3^R\hat{v}_3^{\text{tot}}(\mathbf{x}^R,s)+s\rho x_3^R\hat{v}_3^H(\mathbf{x}^R,s)\\
&\quad=-s\rho_{\text{sens}}x_3^R\Bigg[\frac{1}{2}+\frac{1}{s\rho_{\text{sens}}x_3^R}\\
&\qquad\times\int_{\mathbf{x}\in\partial D_{\text{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)}\hat{G}^H(\mathbf{x}^R|\mathbf{x},s)n_3\,dA\\
&\qquad+\frac{1}{x_3^R}\int_{\mathbf{x}\in\partial D_{\text{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)}x_3\hat{\Gamma}_l^H(\mathbf{x}^R|\mathbf{x},s)n_l\,dA\Bigg]\hat{v}_3^{\text{tot}}(\mathbf{x}^R,s)\\
&\quad=-s\rho_{\text{sens}}x_3^RB\hat{v}_3^{\text{tot}}(\mathbf{x}^R,s),
\end{aligned}
\tag{26}
$$

where the factor B was as defined before [Eq. (23)]. So the interaction coupling coefficient C_I for velocity-sensitive sensors can be determined as

$$
C_I=\frac{\hat{v}_3^{\text{tot}}}{\hat{v}_3^H}=\frac{\rho}{\rho_{\text{sens}}}\frac{1}{1-B}. \tag{27}
$$

F. Cylindrical contrast

The above expression for the coupling coefficient includes the factor B, which is geometry-dependent. In this section the object is considered as a cylinder, since it resembles the spike of a geophone, which is a velocity-sensitive device. For such an object some of the integrals involved can be determined analytically.

The origin of the coordinate system is chosen as the midpoint of the top of the cylinder. The configuration is given in Fig. 1. The scattering surface ∂D_{sct} is divided into three parts: the top (∂D_{top}), mantle (∂D_{man}), and bottom of the cylinder (∂D_{bot}).

First, the first integral in the expression for B is considered. First, n_3 is zero, thus the product $\hat{v}_3^{\text{tot}}n_3$ is zero at the mantle of the cylinder. Second, the Green's function $\hat{G}^H(\mathbf{x}^R|\mathbf{x},s)$ vanishes at the surface $x_3=0$, which follows from Eq. (A1). The top of the cylinder is assumed to be at the free surface, so this holds for the integral over ∂D_{top}. It follows that

$$
\int_{\mathbf{x}\in\partial D_{\text{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)}\hat{G}^H(\mathbf{x}^R|\mathbf{x},s)n_3\,dA=\int_{\mathbf{x}\in\partial D_{\text{bot}}}\hat{G}^H(\mathbf{x}^R|\mathbf{x},s)dA. \tag{28}
$$

Using Eq. (A1), and considering that at ∂D_{bot}, $x_3=h$, the integral over the bottom of the cylinder can be evaluated analytically:

$$
\begin{aligned}
&\int_{\mathbf{x}\in\partial D_{\text{bot}}\setminus(bfx=\mathbf{x}^R)}\hat{G}^H(\mathbf{x}^R|\mathbf{x},s)dA\\
&\quad=\int_{\theta=0}^{2\pi}\int_{r=0^+}^{R}\frac{s\rho}{4\pi}\Bigg[\frac{\exp\left(-\frac{s}{c}r\right)}{r}\\
&\qquad-\frac{\exp\left(-\frac{s}{c}\sqrt{r^2+4h^2}\right)}{\sqrt{r^2+4h^2}}\Bigg]r\,dr\,d\theta\\
&\quad=\frac{\rho c}{2}\Bigg[-\exp\left(-\frac{s}{c}R\right)+\exp\left(-\frac{s}{c}\sqrt{R^2+4h^2}\right)\\
&\qquad+1-\exp\left(-2\frac{s}{c}h\right)\Bigg].
\end{aligned}
\tag{29}
$$

Next, consider the second integral in the expression for B. Since Γ_3^H is bounded at ∂D_{top}, the integral over the top of the cylinder vanishes since x_3 is zero at the top. At the bottom of the cylinder, $\mathbf{n}$ is given by (0,0,1), and x_3 is constant ($=h$). On the mantle of the cylinder, $\hat{\Gamma}_l^Hn_l$ is given by $\hat{\Gamma}_l^H\cos\theta+\hat{\Gamma}_2^H\sin\theta$. Because of these considerations,

$$
\begin{aligned}
&\int_{\mathbf{x}\in\partial D_{\text{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)}x_3\hat{\Gamma}_l^H(\mathbf{x}^R|\mathbf{x},s)n_l\,dA\\
&\quad=\int_{\mathbf{x}\in\partial D_{\text{man}}}x_3(\hat{\Gamma}_l^H\cos\theta+\hat{\Gamma}_2^H\sin\theta)dA\\
&\qquad+h\int_{\mathbf{x}\in\partial D_{\text{bot}}\setminus(\mathbf{x}-\mathbf{x}^R)}\hat{\Gamma}_3^H(\mathbf{x}^R|\mathbf{x},s)dA.
\end{aligned}
\tag{30}
$$

Now, the integral over the mantle of the cylinder can be written as

$$
\begin{aligned}
\int_{\mathbf{x}\in\partial D_{\text{man}}}x_3\hat{\Gamma}_l^H(\mathbf{x}^R|\mathbf{x},s)n_l\,dA=&-\frac{R^2}{4\pi}\int_{\theta=0}^{2\pi}\int_{x_3=0^+}^{h}x_3\Bigg\{\frac{s}{c}\Bigg[\frac{\exp\left(-\frac{s}{c}\sqrt{R^2+(x_3-h)^2}\right)}{R^2+(x_3-h)^2}-\frac{\exp\left(-\frac{s}{c}\sqrt{R^2+(x_3+h)^2}\right)}{R^2+(x_3+h)^2}\Bigg]\\
&+\frac{\exp\left(-\frac{s}{c}\sqrt{R^2+(x_3-h)^2}\right)}{(R^2+(x_3-h)^2)^{3/2}}-\frac{\exp\left(-\frac{s}{c}\sqrt{R^2+(x_3+h)^2}\right)}{(R^2+(x_3+h)^2)^{3/2}}\Bigg\}dx_3\,d\theta.
\end{aligned}
\tag{31}
$$

Using Eq. (A4), the integral over the bottom of the cylinder can be evaluated analytically:

$$h\int_{\mathbf{x}\in\partial D_{\text{bot}}\setminus(\mathbf{x}-bfx^R)}\hat{\Gamma}_3^H(\mathbf{x}^R|\mathbf{x},s)dA=\frac{h^2}{2\pi}\int_{0=0}^{2\pi}\int_{r=0^+}^{R}\exp\left(-\frac{s}{c}\sqrt{r^2+4h^2}\right)\left[\frac{s}{c}\frac{1}{(r^2+4h^2)}+\frac{1}{(r^2+4h^2)^{3/2}}\right]rd$$

$$=h\frac{\exp\left(-2\frac{s}{c}h\right)}{2}-\frac{h^2}{\sqrt{R^2+4h^2}}\exp\left(-\frac{s}{c}\sqrt{R^2+4h^2}\right). \tag{32}$$

The total expression for B becomes:

$$B=\frac{1}{2}-\frac{1}{s\rho_{\text{sens}}h}\frac{\rho c}{2}\left[-\exp\left(-\frac{s}{c}R\right)+\exp\left(-\frac{s}{c}\sqrt{R^2+4h^2}\right)+1-\exp\left(-2\frac{s}{c}h\right)\right]$$
$$+\frac{\exp\left(-2\frac{s}{c}h\right)}{2}-\frac{h\exp\left(-\frac{s}{c}\sqrt{R^2+4h^2}\right)}{\sqrt{R^2+4h^2}}-\frac{R^2}{2h}\int_{x_3=0}^{h}x_3\left\{\frac{s}{c}\left[\frac{\exp\left(-\frac{s}{c}\sqrt{R^2+(x_3-h)^2}\right)}{R^2+(x_3-h)^2}\right.\right.$$
$$\left.\left.-\frac{\exp\left(-\frac{s}{c}\sqrt{R^2+(x_3+h)^2}\right)}{R^2+(x_3+h)^2}\right]+\frac{\exp\left(-\frac{s}{c}\sqrt{R^2+(x_3-h)^2}\right)}{(R^2+(x_3-h)^2)^{3/2}}-\frac{\exp\left(-\frac{s}{c}\sqrt{R^2+(x_3+h)^2}\right)}{(R^2+(x_3+h)^2)^{3/2}}\right\}dx_3. \tag{33}$$

G. Analytical expression for low frequencies for cylindrical contrast

When the arguments of the exponential functions are smaller than 1, the exponential functions can be approximated by the first few terms of its Taylor expansions. The largest term occurring in the exponential functions is $(s/c)\sqrt{R^2+4h^2}$, so the approximation for the whole expression for B holds for $s<c/\sqrt{R^2+4h^2}$, which means small spikes, or equivalently, low frequencies.

In the expression for B, terms are taken only up to the order s^1. The integrals in B can be determined analytically, using the tables from Gradshteyn and Ryzhik (1980, Tables 2.214.2, 2.145.2, 2.271.4, 2.271.5, 2.271.7):

$$\int_{x_3=0}^{h}x_3\left\{\frac{\exp\left(-\frac{s}{c}\sqrt{R^2+(x_3-h)^2}\right)}{R^2+(x_3-h)^2}-\frac{\exp\left(-\frac{s}{c}\sqrt{R^2+(x_3+h)^2}\right)}{R^2+(x_3+h)^2}\right\}dx_3\approx\frac{1}{2}\ln\left(\frac{R^2}{R^2+4h^2}\right)+\frac{h}{R}\arctan\left(\frac{2h}{R}\right), \tag{34}$$

$$\int_{x_3=0}^{h}x_3\left\{\frac{\exp\left(-\frac{s}{c}\sqrt{R^2+(x_3-h)^2}\right)}{(R^2+(x_3-h)^2)^{3/2}}-\frac{\exp\left(-\frac{s}{c}\sqrt{R^2+(x_3+h)^2}\right)}{(R^2+(x_3+h)^2)^{3/2}}\right\}dx_3$$
$$\approx\left[-\frac{1}{R}+\frac{1}{\sqrt{R^2+4h^2}}+\frac{2h^2}{R^2\sqrt{R^2+4h^2}}\right]+\frac{s}{c}\left[\frac{1}{2}\ln\left(\frac{R^2+4h^2}{R^2}\right)-\frac{h}{R}\arctan\left(\frac{2h}{R}\right)\right]. \tag{35}$$

Using these approximations, the expression for the coupling coefficient for motion-sensitive sensors is obtained:

$$C_I=\frac{\rho 2h\sqrt{R^2+4h^2}}{\rho_{\text{sens}}(R^2+4h^2-R\sqrt{R^2+4h^2})-\rho(R^2+4h^2-(R+2h)\sqrt{R^2+4h^2})}. \tag{36}$$

What can be remarked here is that C_I does not contain any phase factor, and not even any velocity c; notice also that C_I is unity when $\rho=\rho_{\text{sens}}$.

H. Block-shaped contrast

This contrast is chosen because, for seismological stations, sometimes concrete blocks are used for mounting. The procedure as used for the cylindrical contrast will be followed. A full expression for the coupling and an analytical approximation for low frequencies will be given.

The configuration is given in Fig. 2. Again, the point of observation is chosen in the midpoint of the top of the block. The block is divided into the top (∂D_{top}), the sides ($\partial D_{\text{sides}}$), and the bottom ($\partial D_{\text{bot}}$).

First, the first integral in the expression for B is considered. As for the cylinder, the contributions for the top and the sides vanish. As a consequence,

$$\int_{\mathbf{x}\in\partial D_{\mathrm{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)} \hat{G}^H(\mathbf{x}^R|\mathbf{x},s)n_3\,dA$$
$$= \int_{\mathbf{x}\in\partial D_{\mathrm{bot}}} \hat{G}^H(\mathbf{x}^R|\mathbf{x},s)dA$$
$$= \frac{\rho c}{2}\left[1-\exp\left(-\frac{s}{c}2h\right)\right]$$
$$+\frac{2\rho c}{\pi}\int_0^{\pi/4}\left[\exp\left(-\frac{s}{c}\sqrt{d^2+d^2\tan^2\theta+4h^2}\right)\right.$$
$$\left.-\exp\left(-\frac{s}{c}\sqrt{d^2+d^2\tan^2\theta}\right)\right]d\theta. \tag{37}$$

Next, the second integral in the expression for B is considered. The contribution from $\partial D_{\mathrm{top}}$ vanishes since x_3 is zero. It follows that

$$\int_{\mathbf{x}\in\partial D_{\mathrm{sct}}\setminus(\mathbf{x}=\mathbf{x}^R)} x_3\hat{\Gamma}_l^H(\mathbf{x}^R|\mathbf{x},s)n_l\,dA$$
$$= \int_{\mathbf{x}\in\partial D_{\mathrm{sides}}} x_3\hat{\Gamma}_l^H n_l\,dA$$
$$+h\int_{\mathbf{x}\in\partial D_{\mathrm{bot}}\setminus(\mathbf{x}=bfx^R)}\hat{\Gamma}_3^H(\mathbf{x}^R|\mathbf{x},s)dA. \tag{38}$$

The integral over the sides all have the same shape, and add up to give

$$\int_{\mathbf{x}\in\partial D_{\mathrm{sides}}} x_3\hat{\Gamma}_l^H(\mathbf{x}^R|\mathbf{x},s)n_l\,dA = -\frac{2d}{\pi}\int_{x_1=0}^{d}\int_{x_3=0}^{h}x_3\left\{\frac{s}{c}\left[\frac{\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3-h)^2}\right)}{r^2+(x_3-h)^2}-\frac{\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3+h)^2}\right)}{r^2+(x_3+h)^2}\right]\right.$$
$$\left.+\frac{\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3-h)^2}\right)}{(r^2+(x_3-h)^2)^{3/2}}-\frac{\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3+h)^2}\right)}{(r^2+(x_3+h)^2)^{3/2}}\right\}dx_3\,dx_1, \tag{39}$$

in which r is given by $r^2=x_1^2+d^2$. The integral over the bottom becomes:

$$h\int_{\mathbf{x}\in\partial D_{\mathrm{bot}}\setminus(\mathbf{x}=\mathbf{x}^R)}\hat{\Gamma}_3^H(\mathbf{x}^R|\mathbf{x},s)dA$$
$$= \frac{4h^2}{\pi}\int_{\theta=0}^{\pi/4}\int_{r=0^+}^{\sqrt{d^2+d^2\tan^2\theta}}\exp\left(-\frac{s}{c}\sqrt{r^2+4h^2}\right)$$
$$\times\left[\frac{s}{c}\frac{1}{r^2+h^2}+\frac{1}{(r^2+4h^2)^{3/2}}\right]r\,dr\,d\theta$$
$$= \frac{h}{2}\exp\left(-\frac{s}{c}2h\right)$$
$$-\frac{4h^2}{\pi}\int_{\theta=0}^{\pi/4}\frac{\exp\left(-\frac{s}{c}\sqrt{d^2+d^2\tan^2\theta+4h^2}\right)}{\sqrt{d^2+d^2\tan^2\theta+4h^2}}d\theta. \tag{40}$$

With these expressions, the exact total coupling coefficient can be determined as for the cylindrical case.

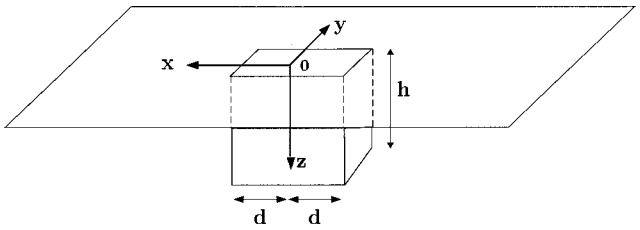


FIG. 2. Configuration for block-shaped contrast.

I. Analytical expression for low frequencies for block-shaped contrast

In the cylindrical case, an analytical expression was derived for low frequencies; in the case of the block, the expressions are more elaborate and therefore the expression has only be determined for $\omega=0$. This means that here it is not shown directly that the first-order term vanishes, although numerically it also showed a negligible small contribution in the first-order (imaginary) term.

The integrals for the block-shaped contrast have again be determined using the tables from Gradshteyn and Ryzhik (1980, Tables 2.271.4, 2.141.2, 2.275.4):

$$\int_{x_1=0}^{d}\int_{z=0}^{h}x_3\left[\frac{1}{(r^2+(x_3-h)^2)^{3/2}}\right.$$
$$\left.-\frac{1}{(r^2+(x_3+h)^2)^{3/2}}\right]dx_1\,dx_3$$
$$= -\ln\frac{(1+\sqrt{2})(d^2+4h^2)^{1/2}}{(d+\sqrt{2d^2+4h^2})}+\frac{h}{d}\arctan\left(\frac{2h}{\sqrt{2d^2+4h^2}}\right) \tag{41}$$

$$\int_0^{\pi/4}[\sqrt{d^2+d^2\tan^2\theta}-\sqrt{d^2+d^2\tan^2\theta+4h^2}]d\theta$$
$$= +d\ln\frac{(\sqrt{d^2+4h^2})}{(d+\sqrt{2d^2+4h^2})}\frac{(\sqrt{2}+1)^{1/2}}{(\sqrt{2}-1)^{1/2}}-h\pi$$
$$+2h\arctan\left(\frac{\sqrt{2d^2+4h^2}}{2h}\right) \tag{42}$$

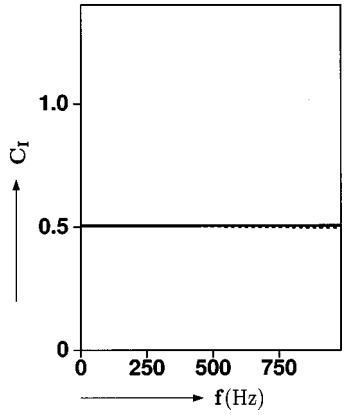


FIG. 3. Velocity-coupling function for cylinder as function of frequency f; $\rho_{\text{sens}}=3000\ \text{kg/m}^3$, $\rho=1500\ \text{kg/m}^3$, $h=0.1$ m, $R=0.005$ m, $c=200$ m/s. Solid line: exact; dashed line: low frequency approximation.

$$\int_0^{\pi/4}\frac{1}{\sqrt{d^2+d^2\tan^2\theta+4h^2}}d\theta = \frac{\pi}{4h}-\frac{1}{2h}\arctan\left(\frac{\sqrt{2d^2+4h^2}}{2h}\right). \tag{43}$$

The total expression for frequency zero then becomes

$$C_I=\frac{\rho}{\rho_{\text{sens}}b_1-\rho b_2}, \tag{44}$$

in which b_1 and b_2 are given by:

$$b_1=1-\frac{2}{\pi}\arctan\left(\frac{d^2}{h\sqrt{2d^2+4h^2}}\right) - \frac{d}{\pi h}\ln\frac{(1+\sqrt{2})^2(d^2+4h^2)}{(d+\sqrt{2d^2+4h^2})^2}, \tag{45}$$

$$b_2=1-\frac{4}{\pi}\arctan\left(\frac{\sqrt{2d^2+4h^2}}{2h}\right) - \frac{d}{\pi h}\ln\frac{(\sqrt{2}+1)}{(\sqrt{2}-1)}\frac{(d^2+4h^2)}{(d+\sqrt{2d^2+4h^2})^2}. \tag{46}$$

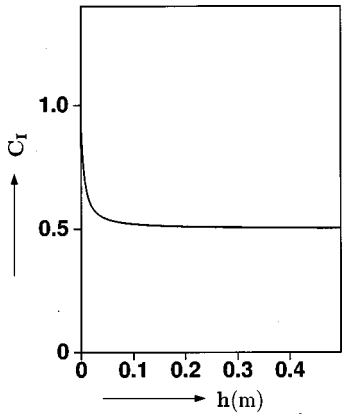


FIG. 4. Velocity-coupling function for cylinder as function of height h; $f=200$ Hz, $\rho_{\text{sens}}=3000\ \text{kg/m}^3$, $\rho=1500\ \text{kg/m}^3$, $R=0.005$ m, $c=200$ m/s. Solid line: exact; dashed line: low frequency approximation.

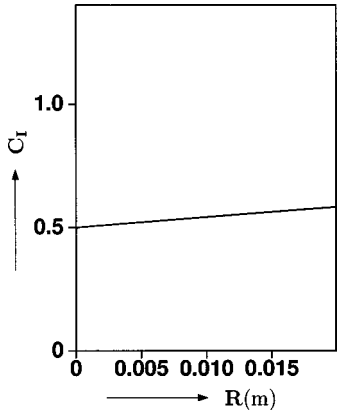


FIG. 5. Velocity-coupling function for cylinder as function of radius R; $f=200$ Hz, $\rho_{\text{sens}}=3000\ \text{kg/m}^3$, $\rho=1500\ \text{kg/m}^3$, $h=0.1$ m, $c=200$ m/s. Solid line: exact; dashed line: low frequency approximation.

II. NUMERICAL RESULTS

In Fig. 3, the velocity-coupling for a cylindrical object is plotted as a function of the frequency. The parameters are chosen for realistic circumstances of geophones as used for land seismics. Both the results of the analytical approximation for low frequencies and the exact representation—the integrals are calculated using a numerical technique—are shown. It is clear that the approximation for low frequencies is useful for these sizes of the sensor, since their values almost coincide on the figures. Also, the coupling is rather insensitive for frequency changes. In this figure, the contrasting domain is rather small. Notice that for the limit to zero frequency, the coupling reaches the value ρ/ρ_{sens}.

In the next two figures, the influence of the geometry of the spike is shown. In Fig. 4 the influence of the height h is shown; the approximation is very good for common values of the geophone dimensions and coincides with the exact value. Notice here that for the limit to zero height, the coupling approaches unity. In Fig. 5, the influence of the radius R on the coupling is shown. Again, the approximation is very good for these parametera and coincides with the exact values; for the limit to zero radius, the coupling approaches ρ/ρ_{sens}.

Geophones are velocity-measuring devices, while hydrophones are pressure-measuring devices. In Fig. 6 the

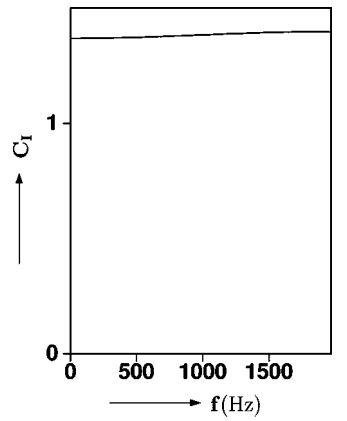


FIG. 6. Pressure-coupling function for cylinder as function of frequency f; $\rho_{\text{sens}}=1500\ \text{kg/m}^3$, $\rho=1000\ \text{kg/m}^3$, $h=0.005$ m, $R=0.025$ m, $c=1500$/s.

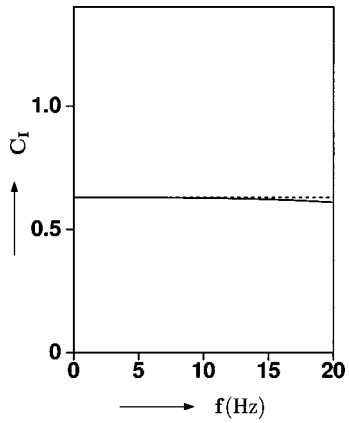


FIG. 7. Velocity-coupling function for block as function of frequency f; $\rho_{sens}=3000\ \mathrm{kg/m^3}$, $\rho=1500\ \mathrm{kg/m^3}$, $h=1$ m, $d=1$ m, $c=200$ m/s. Solid line: exact; dashed line: low frequency approximation.

pressure-coupling for a hydrophone is shown as a function of the frequency. It turns out that this coupling is not very dependent on the frequency. In the next figures, the velocity-coupling for a blocked shape with larger dimensions and using lower frequencies is calculated. The values here are chosen for realistic circumstances of seismological stations as used in seismology. From Fig. 7, it follows that the low frequency approximation is very good again for these sizes of the block. Again, the coupling does not vary much as a function of the frequency. In Figs. 8 and 9, the influence of the geometry is shown. The same behavior as for the cylinder is observed. Geometry factors do have a larger influence on large objects than on small objects. A better coupling is achieved when using thin blocks with large horizontal dimensions.

III. DISCUSSION

In the results, it can be seen that the dominant effect is the density effect; if there is no density contrast, the coupling coefficient is unity. However, it should be realized that the density contrast has arisen because the point of observation is *on* the scattering object and the density has been chosen to be the density of the sensor [see Eq. (17)]. It is arguable whether this density should be the average between the density of the sensor and that of the surrounding medium. In that case, the coupling coefficient would be different in their density term.

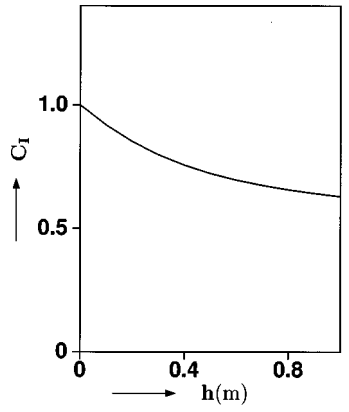


FIG. 8. Velocity-coupling function for block as function of height h; $f=2.5$ Hz, $\rho_{sens}=3000\ \mathrm{kg/m^3}$, $\rho=1500\ \mathrm{kg/m^3}$, $d=1$ m, $c=200$ m/s.

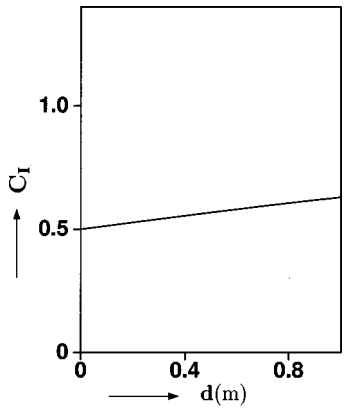


FIG. 9. Velocity-coupling function for block as function of width d; $f=2.5$ Hz, $\rho_{sens}=3000\ \mathrm{kg/m^3}$, $\rho=1500\ \mathrm{kg/m^3}$, $h=1$ m, $c=200$ m/s.

The approach in this paper is that a welded contact between the sensor and the surrounding medium is assumed. In practice, it often occurs that the coupling is not good, and then the model developed here does not account for it. One approach would be to introduce slip along the sensor, by introducing a linear-slip model along the sensor, or by introducing a small ''boundary layer'' which has damping effects. This will be part of future studies.

IV. CONCLUSIONS

In this paper, a general approach for describing the disturbance of a wave field due to the presence of a measuring device has been developed. It was shown that, even when there is perfect material contact between the device and its environment, the measured wave field differs from the wave field that would exists if no measuring device was present. This effect is called the interaction coupling.

The interaction coupling is mainly dependent on the densities of the materials of the measuring device and the surroundings, and on the geometric dimensions of the device. The frequency spectrum of the waves is only of minor importance, as was shown via the low frequency approximations. These low frequency approximations clearly show the main physics of interaction coupling.

APPENDIX: GREEN'S FUNCTIONS FOR CYLINDRICAL OBJECT

In this Appendix explicit expressions are given for the Green's functions $\hat{G}^H$ and $\hat{\Gamma}_k^H$ for the half-space which are needed for evaluating the interaction coupling. The contrasting domain is considered as a cylinder, so it is advantageous to turn to cylindrical coordinates. The cylinder is vertically planted and with its upper plane equal to the surface of the earth. The middle of the cylinder's upper plane is taken to be the origin of a cylindrical coordinate system. The x_3-axis is pointing downward. Use will be made of Eq. (16), where the fixed point x^R has to be located on the boundary of the cylinder. For this reason, the fixed point is taken to be the middle point of the cylinder's lower plane.

The Green's function $\hat{G}^H$ becomes:

$$\hat{G}^H(\mathbf{x}^R|\mathbf{x},s)=\frac{s\rho}{4\pi}\left[\frac{\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3-h)^2}\right)}{\sqrt{r^2+(x_3-h)^2}}-\frac{\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3+h)^2}\right)}{\sqrt{r^2+(x_3+h)^2}}\right], \tag{A1}$$

while the Green's function $\hat{\Gamma}_k^H$ is given by

$$\hat{\Gamma}_1^H(\mathbf{x}^R|\mathbf{x},s)=-\frac{r\cos\theta}{4\pi}\left\{\frac{s}{c}\left[\frac{\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3-h)^2}\right)}{r^2+(x_3-h)^2}-\frac{\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3+h)^2}\right)}{r^2+(x_3+h)^2}\right]\right.$$
$$\left.+\frac{\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3-h)^2}\right)}{(r^2+(x_3-h)^2)^{3/2}}-\frac{\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3+h)^2}\right)}{(r^2+(x_3+h)^2)^{3/2}}\right\}, \tag{A2}$$

$$\hat{\Gamma}_2^H(\mathbf{x}^R|\mathbf{x},s)=-\frac{r\sin\theta}{4\pi}\left\{\frac{s}{c}\left[\frac{\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3-h)^2}\right)}{r^2+(x_3-h)^2}-\frac{\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3+h)^2}\right)}{r^2+(x_3+h)^2}\right]\right.$$
$$\left.+\frac{\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3-h)^2}\right)}{(r^2+(x_3-h)^2)^{3/2}}-\frac{\exp\left(-\frac{s}{c}\sqrt{x^2+(x_3+h)^2}\right)}{(r^2+(x_3+h)^2)^{3/2}}\right\}, \tag{A3}$$

$$\hat{\Gamma}_3^H(\mathbf{x}^R|\mathbf{x},s)=-\frac{1}{4\pi}\left\{\frac{s}{c}\left[\frac{(x_3-h)\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3-h)^2}\right)}{r^2+(x_3-h)^2}-\frac{(x_3+h)\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3+h)^2}\right)}{r^2+(x_3+h)^2}\right]\right.$$
$$\left.+\frac{(x_3-h)\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3-h)^2}\right)}{(r^2+(x_3-h)^2)^{3/2}}-\frac{(x_3+h)\exp\left(-\frac{s}{c}\sqrt{r^2+(x_3+h)^2}\right)}{(r^2+(x_3+h)^2)^{3/2}}\right\}, \tag{A4}$$

where the following relation has been used:

$$\hat{\Gamma}_k^H(x^R|x,s)=\frac{1}{s\rho}\,\partial_k\hat{G}^H(x^R|x,s). \tag{A5}$$

Fokkema, J. T., and van den Berg, P. M. (**1993**). *Seismic Applications of the Acoustic Reciprocity Theorem* (Elsevier, Amsterdam).

Gradshteyn, I. S., and Ryzhik, I. M. (**1980**). *Table of Integrals, Series and Products* (Academic, Orlando, FL).

Tan, T. H. (**1987**). ''Reciprocity theorem applied to the geophone-ground coupling problem,'' Geophysics **52**, 1715–1717.

Synthesis of optimal, single-frequency, passive control laws, with application to reducing the acoustic radiation from a submerged spherical shell

Paul J. Titterton, Jr.
Applied Physical Sciences Laboratory, SRI International, 333 Ravenswood Avenue, Menlo Park, California 94025 and Space Telecommunications and Radioscience Laboratory, Stanford University, Stanford, California 94305

(Received 5 January 1999; accepted for publication 7 January 1999)

This paper uses (i) a network model for vibrating and radiating structures, (ii) network-scattering variables, and (iii) semi-definite programing to synthesize load admittances which achieve performance bounds for single-frequency, multi-port passive control of acoustic radiation. The network model relates the disturbance, regulated output, and the force and velocity at control-load attachment points. The semi-definite program minimizes the maximum singular value of a weighted, single-frequency, closed-loop, disturbance-to-regulated-output transfer matrix, while maintaining the passivity of the load. Three design examples demonstrate control of single-frequency radiated power from a (mathematically modeled) submerged spherical shell. First, we minimize the power radiated into a 30° far-field sector for two configurations: (i) one disturbance with one control load, and (ii) two disturbances with three control loads. For $4 \leqslant ka \leqslant 7$, radiated power is reduced by 10–20 dB. This reduced radiated power is compared to the conjugate-matched-load radiated power, which is greater than the uncontrolled at almost every frequency. Second, we show that the optimal-passive radiated power is sensitive to the phase of the uncontrolled response at the attachment point of the load. Third, we remove the passivity constraint to show that the best (active or passive) controller is often passive. When the optimal is active, the radiated power is only a few decibels less than the optimal passive. Our technique's primary shortcoming is that the quieting is optimal at only a single frequency. Broadening the control bandwidth is a topic-of current research.
© *1999 Acoustical Society of America.* [S0001-4966(99)01804-4]

PACS numbers: 43.40.Dx [PJR]

INTRODUCTION AND BACKGROUND

In this paper, we use (i) a network model for vibrating and radiating structures, (ii) network-scattering variables, and (iii) semi-definite programing to bound the single-frequency quieting that can be achieved by a passive controller. We also synthesize the passive control law which realizes that bound. The utility of this synthesis algorithm is illustrated by designing several passive controllers which minimize the worst-case power radiated from a submerged spherical shell.

Our control-law synthesis algorithm requires a generalized network model of the structure that we wish to quiet. We refer to the structure itself as *the plant* and to the network model as *the plant model*. The plant model describes the single-frequency relationship between the disturbance sources, the regulated output, and the force and velocity at the controller-attachment points. The disturbance sources, which cause the unwanted noise and vibration, can be physically attached to the structure or physically separated from the structure. The regulated output, which is the noise or vibration that we want to reduce, is measured at some number of discrete locations. Each controller-attachment point is represented as a single port in the plant model. The passive controller, which loads the plant at the attachment points, is described by its admittance matrix or its network-reflection matrix.

Given a plant model, our synthesis algorithm searches over all passive control loads to minimize the maximum singular value of the weighted, single-frequency, closed-loop, disturbance-to-regulated-output transfer matrix. (The maximum singular value of the weighted disturbance-to-regulated-output transfer matrix is the worst-case weighted 2-norm of the regulated output, given a unit-2-norm disturbance.) The controller's passivity is enforced by requiring the average power to flow *from* the vibrating structure *into* the control load. If this necessary condition for passivity is met, we can always realize the single-frequency control law as a network of physically passive elements. Searching over all passive loads to minimize this maximum singular value is readily written as a constrained optimization problem.

Our synthesis algorithm, which is presented in Sec. I, proceeds in the following three steps:

(1) Develop a single-frequency plant model, including disturbance sources, regulated outputs, and controller-attachment points, and define the constrained optimization problem in terms of the plant model and the control load.
(2) Transform the attachment-point force and velocity to the network-scattering variables.
(3) Translate the optimization problem into a semi-definite

program (SDP) and solve for the network-reflection matrix of the control load.

Recent advances in interior-point methods for convex optimization enable fast and accurate solution of SDPs.[1,2]

A control law obtained by this procedure is optimal at only a single frequency. If we were to synthesize optimal control laws at two (or more) adjacent frequencies, there is no guarantee that a network of physically passive elements could realize the optimal load at those frequencies while providing close-to-optimal interpolated values. Broadening the control band is a topic of current research.

In Sec. II, we illustrate the utility of this procedure by designing optimal passive control laws which reduce the acoustic radiation from a mathematically modeled submerged spherical shell. We weight the closed-loop disturbance-to-regulated-output transfer matrix so that the SDP minimizes the acoustic power radiated into a 30° far-field sector. At most frequencies in the band $4 \leqslant ka \leqslant 7$, the optimal control load reduces radiated power by 10–20 dB. We do not discuss the mechanical (or electro-mechanical) design necessary to physically realize the optimal control load.[3]

Passive control offers two important features which are unavailable in active control. First, if the plant is also passive, then the controlled system will remain stable for arbitrary plant variations. Second, the controller can be realized in a manner that requires no expensive processor and no line power, which is particularly useful in low-cost systems and systems which are fielded in remote locations.

In passive noise and vibration control, heuristic methods are ubiquitous. Two well-known examples are provided by Beranek[4] and Timoshenko and Young.[5] Beranek describes a method for reducing broadband radiation from vibrating machinery: the design engineer applies a sequence of vibration-isolation mounts, rigid enclosures, and acoustic absorbing material until the radiated noise reaches a suitable level. Timoshenko and Young describe a method for reducing the vibration caused by single-frequency disturbances:[6] The design engineer places a one-spring–one-mass oscillator—a ''dynamic absorber''—on the vibrating structure and modifies the mass and spring constant to null the disturbance.

Heuristically designed passive controllers are sometimes used to condition a plant prior to active controller design. This use is a result of the rule of thumb which states that a well-damped structure is easier to control. In particular, passive controllers are used as the ''low-authority'' in high-authority, low-authority control.[7,8]

One obstacle to the general application of passive control is the scarcity of synthesis methods which produce optimal-passive control laws. Where synthesis of active control laws allows a variety of optimality criteria and design tools, synthesis of passive control laws does not. Of course, there are exceptions for particular design criteria: for a single-mode resonant system, Timoshenko and Young[5] also describe a method for minimizing the resonant system's maximum displacement by applying a specially designed spring-mass-dash-pot damper; Hagood and von Flowtow[9] extend this single-mode method to electrically terminated piezoelectrics; MacMartin and Hall[10] present an H^{∞} technique for broadband conjugate matching; and Benhabib *et al.*,[11] Lozano-Leal and Joshi,[12] and McLaren and Slater[13] derive a variety of linear quadratic techniques to control vibration in large space structures.

These previously published methods for optimal synthesis of passive control laws restrict the quieting measure and the allowable disturbance-to-regulated-output transfer matrices. Additionally, these methods do not search over all possible passive control laws to achieve their optimum.

In contrast, the single-frequency algorithm presented in this paper searches over all possible passive control laws to minimize a worst-case quadratic function of the regulated-output variables.

This paper is the third in a series which investigates single-frequency circuit methods for acoustic radiation and scattering. The first paper[14] analyzes the scattering from an internally loaded spherical shell. The internal structure loads a 2-port description of the dynamics of the fluid-loaded shell. We apply circuit techniques to interpret the many resonances and to bound the scattering for any passive internal structure. The second paper[15] presents a scattering formulation in which the entire target loads an N-port description of the fluid. This paper investigates the synthesis of passive loads to achieve a desired dynamic response.

The body of this paper is divided into two sections. Section I develops the synthesis algorithm that finds the optimal passive load, and Sec. II provides an example application which illustrates the reduction of acoustic radiation from a submerged spherical shell.

I. THEORETICAL DEVELOPMENT

A. Single-frequency network models for the plant and the controller

Our algorithm requires a network model which describes the single-frequency plant dynamics:

$$\begin{bmatrix} \mathbf{r} \\ \mathbf{v}_l \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{Y}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{d} \\ \mathbf{f}_l \end{bmatrix}. \tag{1}$$

Each element of the vector $\mathbf{d}$ is the complex amplitude of a disturbance source, and each element of $\mathbf{f}_l$ is the complex amplitude of the control force applied at one of the load points. The elements of $\mathbf{v}_l$ are the corresponding load-point velocities, and the elements of $\mathbf{r}$ are measurements of any field variable at the regulated-output locations. We require that $\mathcal{P}_l = \mathrm{Re}\{\mathbf{f}_l^H \mathbf{v}_l\}/2$ is the average power flowing into the structure through the load points. The 2×2-block matrix describes the relationship between $\mathbf{d}$, $\mathbf{f}_l$, $\mathbf{v}_l$, and $\mathbf{r}$. The ith column of $\mathbf{H}_{11}$ is the regulated output for an $\mathbf{e}_i$ disturbance and zero control force ($\mathbf{e}_i$ is a vector of zeros with a one in the ith element); the ith column of $\mathbf{H}_{12}$ is the regulated output for an $\mathbf{e}_i$ control force and zero disturbance; the ith column of $\mathbf{H}_{21}$ is the velocity at the load points for an $\mathbf{e}_i$ disturbance and zero control force; and the ith column of $\mathbf{Y}_{22}$ is the velocity at the load points for an $\mathbf{e}_i$ control force and zero disturbance. We assume that the plant, as seen from the load points, is passive, which implies that $\mathbf{Y}_{22}$ is symmetric with

real part positive semi-definite. In the velocity-current force-voltage analogy, $\mathbf{Y}_{22}$ is clearly an admittance matrix. The other terms are hybrids, used to fill out the linear relationship between field variables. The disturbances can be sources of any type. In this paper's design example, we use force sources, but we could just as easily have chosen velocity sources or a combination of force and velocity sources.

The control load, which is our controller design variable, relates $\mathbf{f}_l$ and $\mathbf{v}_l$. In the admittance representation

$$\mathbf{v}_l = -\mathbf{Y}_l \mathbf{f}_l . \tag{2}$$

We assume that $\mathbf{Y}_l$ is symmetric with real part positive semi-definite, which is the necessary condition for control-load passivity.[3] Direct substitution into Eq. (1) yields the following relationship between the disturbance and regulated output:

$$\mathbf{r} = [\mathbf{H}_{11} - \mathbf{H}_{12}(\mathbf{Y}_{22} + \mathbf{Y}_l)^{-1}\mathbf{H}_{21}]\mathbf{d}, \tag{3}$$

where, of course, the single-frequency, closed-loop, disturbance-to-regulated-output transfer matrix is $[\mathbf{H}_{11} - \mathbf{H}_{12}(\mathbf{Y}_{22} + \mathbf{Y}_l)^{-1}\mathbf{H}_{21}]$.

Now that we have written the disturbance-to-regulated-output transfer matrix as a function of $\mathbf{Y}_l$, and we have a constraint on $\mathbf{Y}_l$ that guarantees passivity, we can write our control-law synthesis problem as the following constrained optimization:

$$\begin{aligned} &\min_{\mathbf{Y}_l = \mathbf{Y}_l^T} \quad \overline{\sigma}\{\mathbf{Q}_1[\mathbf{H}_{11} - \mathbf{H}_{12}(\mathbf{Y}_{22} + \mathbf{Y}_l)^{-1}\mathbf{H}_{21}]\mathbf{Q}_2\} \\ &\text{such that} \quad \mathrm{Re}\{\mathbf{Y}_l\} \geqslant 0. \end{aligned} \tag{4}$$

We minimize the maximum singular value of the weighted, closed-loop, disturbance-to-regulated-output transfer matrix while maintaining passivity of the controller. $\mathbf{Q}_1$ is chosen to emphasize particular combinations of regulated outputs, $\mathbf{Q}_2$ is chosen to emphasize particular combinations of disturbances, and $\overline{\sigma}(\cdot)$ is the maximum singular value of the argument.

B. Transformation to network-scattering variables

The minimization criterion in Eq. (4), as a function of $\mathbf{Y}_l$, is not suitable for optimization; in this section we introduce the network-scattering variables,[3] which transform Eq. (4) into an optimization problem that is more easily solved. The following change of field-variable basis at the load points transforms the admittance representation to the network-scattering representation:[15]

$$\begin{bmatrix} \mathbf{f}_l \\ \mathbf{v}_l \end{bmatrix} = \begin{bmatrix} -\mathrm{Re}\{\mathbf{Y}_{22}\}^{-1/2} & \mathrm{Re}\{\mathbf{Y}_{22}\}^{-1/2} \\ \mathbf{Y}_{22}^H \mathrm{Re}\{\mathbf{Y}_{22}\}^{-1/2} & \mathbf{Y}_{22}\,\mathrm{Re}\{\mathbf{Y}_{22}\}^{-1/2} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}. \tag{5}$$

Substituting Eq. (5) into (1) and (2) yields

$$\begin{bmatrix} \mathbf{r} \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{d} \\ \mathbf{b} \end{bmatrix}; \quad \mathbf{b} = \mathbf{\Gamma}\mathbf{a}, \tag{6}$$

where the $\mathbf{S}_{ij}$ matrices represent the plant dynamics in the new basis, and $\mathbf{\Gamma}$ is the network-reflection matrix of the control load.

The network-reflection matrix is the new design variable. The symmetry of $\mathbf{Y}_l$ and $\mathbf{Y}_{22}$ translates into symmetry of $\mathbf{\Gamma}$; the matrix-positivity condition on $\mathrm{Re}\{\mathbf{Y}_l\}$ translates into a matrix-positivity condition on $[\mathbf{I} - \mathbf{\Gamma}^H\mathbf{\Gamma}]$; and the closed-loop disturbance-to-regulated-output transfer matrix translates into an offset-linear function of the design variable, $[\mathbf{S}_{11} + \mathbf{S}_{12}\mathbf{\Gamma}\mathbf{S}_{21}]$.

Writing the constrained optimization problem in terms of the $\mathbf{S}_{ij}$ and $\mathbf{\Gamma}$ matrices yields

$$\begin{aligned} &\min_{\Gamma = \Gamma^T} \quad \overline{\sigma}\{\mathbf{Q}_1[\mathbf{S}_{11} + \mathbf{S}_{12}\mathbf{\Gamma}\mathbf{S}_{21}]\mathbf{Q}_2\} \\ &\text{such that} \quad \mathbf{I} - \mathbf{\Gamma}^H\mathbf{\Gamma} \geqslant 0, \end{aligned} \tag{7}$$

which, at first glance, does not appear any simpler than Eq. (4).

C. Rewriting the constrained optimization problem as a semi-definite program

Using methods described by Boyd and El Ghaoui[16] and by Vandenberghe and Boyd,[2] we can rewrite Eq. (7) as the following SDP:

$$\begin{aligned} &\min_{\lambda, \Gamma = \Gamma^T} \quad \lambda \\ &\text{such that} \quad \begin{bmatrix} \lambda\mathbf{I} & \mathbf{X}^H & 0 & 0 \\ \mathbf{X} & \lambda\mathbf{I} & 0 & 0 \\ 0 & 0 & \mathbf{I} & \mathbf{\Gamma}^H \\ 0 & 0 & \mathbf{\Gamma} & \mathbf{I} \end{bmatrix} \geqslant 0, \end{aligned} \tag{8}$$

where $\mathbf{X} = \mathbf{Q}_1[\mathbf{S}_{11} + \mathbf{S}_{12}\mathbf{\Gamma}\mathbf{S}_{21}]\mathbf{Q}_2$, and λ is an auxiliary variable.

Equation (8) is a special case of the general SDP, which can be solved using established methods.[17] The general SDP is defined:

$$\begin{aligned} &\min_{\mathbf{x}} \quad \mathbf{c}^T\mathbf{x} \\ &\text{such that} \quad F(\mathbf{x}) = F_0 + F_1 x_1 + \cdots + F_n x_n \geqslant 0, \end{aligned} \tag{9}$$

where $\mathbf{x} = [x_1, \cdots, x_n]^T$, is a vector of design variables, $\mathbf{c}$ is a weighting vector used to define the direction of minimization, and F is a real, symmetric, affine matrix function of $\mathbf{x}$. A computer program for solving SDPs is available by anonymous FTP from is1.stanford.edu/pub/boyd. This concludes our development of a synthesis algorithm which produces optimal-passive control laws.

II. SINGLE-FREQUENCY LIMITS ON PASSIVE CONTROL OF ACOUSTIC RADIATION FROM A SUBMERGED SPHERICAL SHELL THAT HAS MULTIPLE DISTURBANCE SOURCES

This section presents several numerical experiments which examine different aspects of the passive control of acoustic radiation from a submerged spherical shell. The

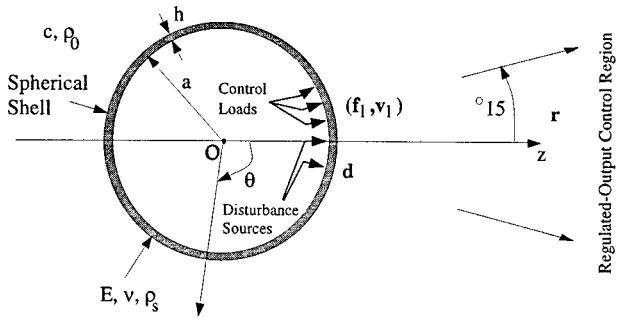


FIG. 1. Physical configuration for passive control of the spherical shell.

physical configuration is shown in Fig. 1, and characteristics of the particular numerical experiments—including disturbance locations, load points, regulated outputs, temporal frequencies, and control-load types—are detailed in Table I. In experiment 1a, the plant has (i) a single disturbance source, located at $\theta=0$, where θ is degrees along the meridian measured from the $+z$ axis, (ii) a single control load, located at $(\theta,\phi)=(17°,175°)$, where ϕ is degrees longitude measured from the $+x$ axis, and (iii) regulated output locations covering the 30° far-field sector centered on $\theta=0$. For the frequency band $4\leqslant ka\leqslant 7$, we compare the radiated power for three different control laws: (i) the optimal-passive load, (ii) the conjugate-matched load, and (iii) no load. The ''no-load'' control law is given by a zero impedance or an infinite admittance—that is, the attachment point force is zero, and the velocity is unconstrained. This, of course, describes a controller that is physically absent. Experiment 1b repeats the comparison of experiment 1a, but uses a plant with two disturbance sources and three control loads. The particular disturbance locations and load points were chosen arbitrarily. Experiment 2 shows how the optimal radiated power changes, sometimes dramatically, with the position of the controller. Experiments 3a and 3b modify the passivity constraint to examine its effect on the optimal radiated power. Experiment 3a finds the optimal active control load by removing the power-flow constraint at the load points. For the frequency band $4\leqslant ka\leqslant 7$, we compare the radiated power for three different control laws: (i) the optimal-passive load, (ii) the optimal-active load, and (iii) no-load. Experiment 3b varies the power-flow constraint to determine the tradeoff between radiated power and the power that the load can supply to the plant.

With reference to Fig. 1, the specifics of the plant configuration are as follows: the disturbances are independent force sources, which are normally applied to the interior of the shell; the control-load forces are also normally applied to the interior of the shell; the control-load velocities are oriented outward; and the regulated output is far-field radiated pressure on the mesh ($\theta_i=1.25,3.75,6.25,\cdots,13.75$; $\phi_j=0,2.5,5,\cdots,357.5$). We refer to the far-field sector of regulated outputs as the control region. The material parameters of the shell and fluid are as follows: the shell's mid-plane radius, $a=5$ m; thickness, $h=0.15$ m; Young's modulus, $E=19.6\times10^{10}\,\mathrm{N/m^2}$; Poisson's ratio, $\nu=0.3$; density, $\rho_s=7668.7\,\mathrm{kg/m^3}$; the speed of sound in the surrounding fluid, $c=1480$ m/s; and the fluid's density, $\rho_0=1000\,\mathrm{kg/m^3}$.

We calculate the plant model using the standard bending shell equations from Junger and Feit (J&F):[18] $\mathbf{H}_{11}$ and $\mathbf{H}_{12}$ are calculated using J&F Eq. (9.14); and $\mathbf{H}_{21}$ and $\mathbf{Y}_{22}$ are calculated using J&F Eq. (9.13). The shell admittance and Helmholtz admittance, which are required to implement Eqs. (9.13) and (9.14), are given by J&F Eqs. (7.121) and (6.29), respectively.

In Eq. (4), the constrained optimization problem, we take $\mathbf{Q}_1$ to be the square root of the radiation admittance matrix. {The acoustic power radiated into our control region is $P_{\mathrm{rad}}=(1/\rho_0 c)\int p(\theta,\phi)^* p(\theta,\phi)\,d\sigma$, where the region of integration is $\theta=[0,15°]$ and $\phi=[0,360°]$, and $p(\theta,\phi)$ is the far-field radiated pressure. In our frequency range, the regulated output—which is point measurements of $p(\theta,\phi)$ on a far-field mesh—can be used as the coefficients of a subsectional basis which represents $p(\theta,\phi)$ over the control region. Substituting this discretization into the above integral yields $P_{\mathrm{rad}}=\mathbf{r}^H\mathbf{Q}_1^H\mathbf{Q}_1\mathbf{r}$, where $\mathbf{Q}_1^H\mathbf{Q}_1$ is the radiation admittance matrix.} We take $\mathbf{Q}_2=\mathbf{I}$. For this choice of $\mathbf{Q}_1$ and $\mathbf{Q}_2$, the numerical value of the cost function is the square root of the

TABLE I. Characteristics of particular numerical experiments on the passive control of acoustic radiation from a submerged spherical shell.

	Dist.		Load				
Exp.	θ	ϕ	θ	ϕ	Reg. out.	Freq.	Control-load type
1a	0		17	175	30° sector	$4\leqslant ka\leqslant 7$	(i) Optimal passive (ii) Conj. matched (iii) No-load
1b	0		17	175	30° sector	$4\leqslant ka\leqslant 7$	(i) Optimal passive
	20	139	18	23			(ii) Conj. matched
			19	72			(iii) No-load
2	0		$10\leqslant\theta\leqslant 40$	175	30° sector	$ka=5.49$	Optimal passive
3a	0		17	175	30° sector	$4\leqslant ka\leqslant 7$	(i) Optimal passive (ii) Optimal active (iii) No-load
3b	0		17	175	30° sector	$ka=5.49$	Varying power-flow constraint

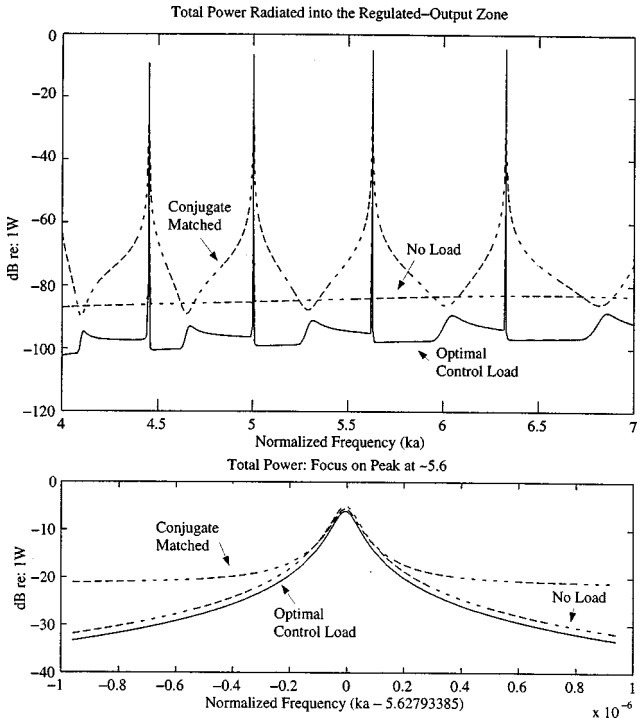


FIG. 2. Numerical experiment 1a: radiated power achieved using the optimal-passive load, the conjugate-matched load, and no load in the single-disturbance single-control-load plant configuration.

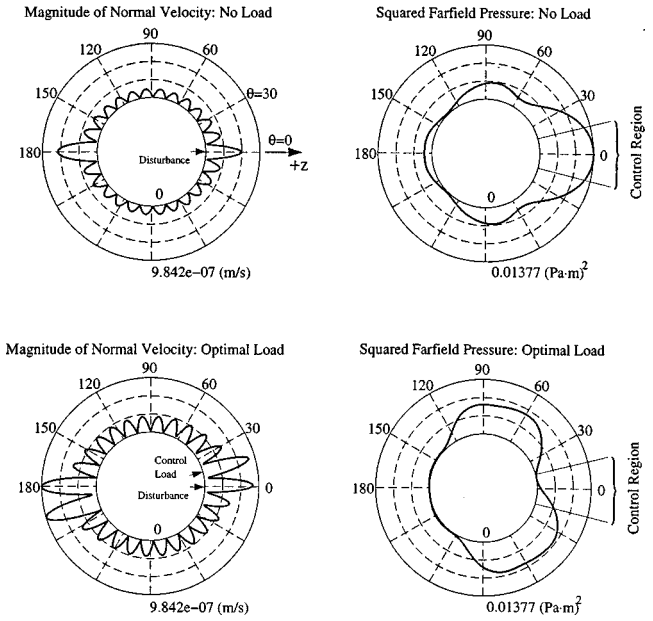


FIG. 3. Numerical experiment 1a: normal velocity magnitude and squared radiated pressure, at $ka=5.9$, using the optimal-passive load and no-load in the single-disturbance single-control-load plant configuration.

total acoustic power radiated into the control region by an evenly weighted set of disturbance sources. Translating the constrained optimization problem to an SDP, and solving—as per Sec. I—minimizes the radiated power and returns the network-reflection matrix of the optimal-passive load.

A. Numerical experiments 1a and 1b: Radiated power achieved with (i) the optimal-passive load, (ii) the conjugate-matched load, and (iii) no-load

In this section, we examine radiated power as a function of frequency, and we examine surface velocity patterns and far-field pressure patterns at a single frequency.

The results of experiment 1a are shown in Fig. 2, which plots the unit-disturbance radiated power for (i) the optimal-passive load, (ii) the conjugate-matched load, $\mathbf{Y}_l=\mathbf{Y}_{22}^H$, and (iii) no load. The upper panel shows power versus frequency over the normalized frequency range $4\leqslant ka\leqslant 7$. The narrow peaks correspond to the lower-branch resonances of the submerged spherical shell.[19] The lower panel focuses on the resonance near $ka=5.6$. At this resonance, the optimal-passive controller is ineffective.

This figure shows that the optimal passive load reduces radiated power by 10–20 dB—except at resonances—while the conjugate-matched load almost always enhances radiated power. The conjugate-matched load, which maximizes the power into the load, is an often-suggested rule of thumb for passive vibration control. To draw maximum power, this load sets up large amplitude waves on the sphere. At most frequencies, these waves increase radiated power.

Additional detail for experiment 1a is provided in Fig. 3: The upper panels show the magnitude of the surface velocity and the squared magnitude of the far-field pressure—at $ka=5.49$—for no, load; the lower panels show the magnitude of the surface velocity and the squared magnitude of the far-field pressure—also at $ka=5.49$—for the optimal-passive load. The surface and far-field quantities are plotted on the $\phi=175°$ great circles. The disturbance and load locations are shown on the velocity plots, and the control region is shown on the far-field pressure plots. The dominant mode in the velocity distribution, $n=14$, has small radiation resistance and does not appear in the radiated pressure. It is interesting to note that the velocity distribution for the optimal-passive load has greater amplitude than that for no load. The greater-amplitude pattern places a null in the control region, but enhances the radiated pressure outside the control region.

In experiment 1b, we repeat these analyses for the two-disturbance three-load-point plant configuration. Figure 4 shows the worst-case radiated power given a unit-norm disturbance for the three different control laws. The upper panel shows power versus frequency for $4\leqslant ka\leqslant 7$, and the lower panel focuses on the resonance near $ka=5.6$. For the larger number of load points, the radiation reduction achieved by the optimal load is greater, and the radiation enhancement by the conjugate-matched load is also greater. Figure 5 gives the surface velocity and far-field pressure patterns. For the larger number of control loads, the velocity distribution for the optimal-passive load is still greater than that for no load. The greater-amplitude pattern reduces the radiated field everywhere inside the control region, but still enhances the radiated field outside the control region.

B. Numerical experiment 2: Effect of control-load location on radiation reduction

In experiment 2, we show how optimal radiated power changes with the position of the control load. A single disturbance is placed at the sphere's north pole, and we calcu-

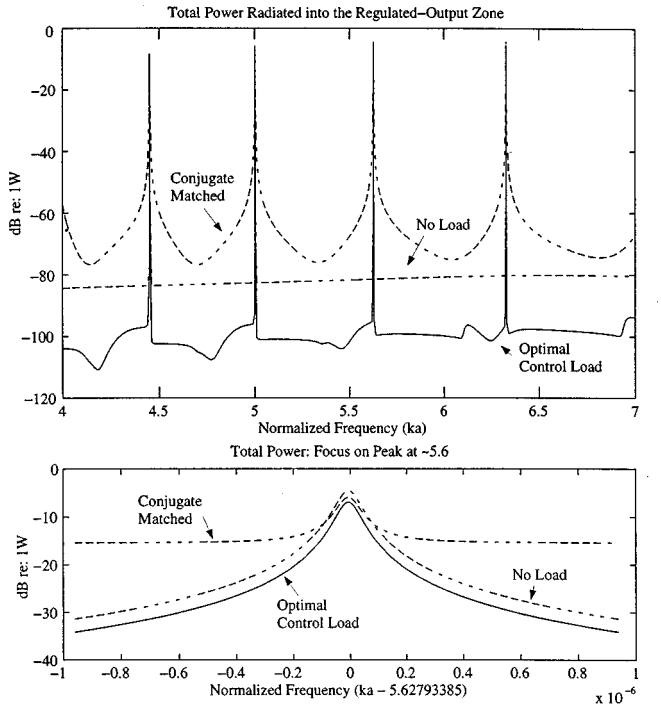


FIG. 4. Numerical experiment 1b: radiated power achieved using the optimal-passive load, the conjugate-matched load, and no-load in the two-disturbance three-control-load plant configuration.

late the optimal load and radiated power as the load's location is moved from $\theta=10°$ to $\theta=40°$ in 0.2° increments. All calculations are performed at $ka=5.49$. Figure 6 plots the optimal radiated power—normalized by the no-load radiated power—as a function of load location. The asterisk indicates the load location in experiment 1a.

This plot shows that loads located close to the disturbance most effectively reduce radiated power; as the load moves away from the disturbance the power reduction tends to decrease. When the control load is close to the disturbance source, the radiation pattern of the controller and the radiation pattern of the disturbance are alike. As the control load moves away from the disturbance source, their patterns become less and less alike.

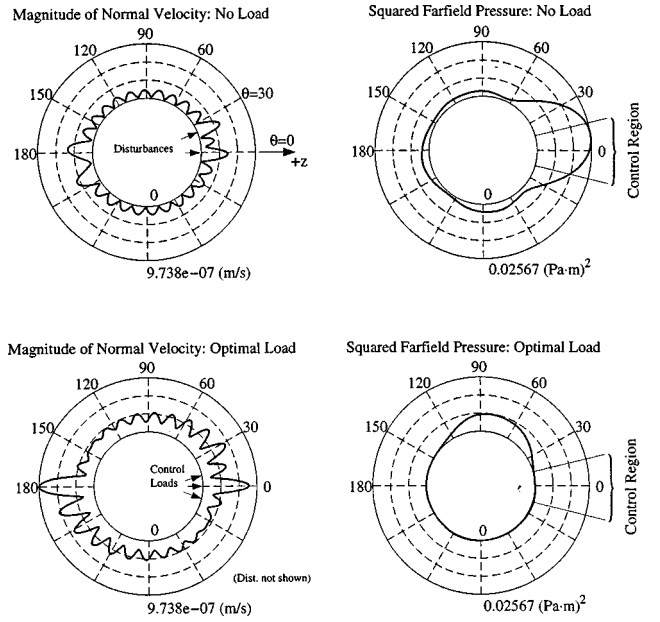


FIG. 5. Numerical experiment 1b: normal velocity and radiated pressure, at $ka=5.49$, using the optimal-passive load and no-load in the two-disturbance three-control-load plant configuration.

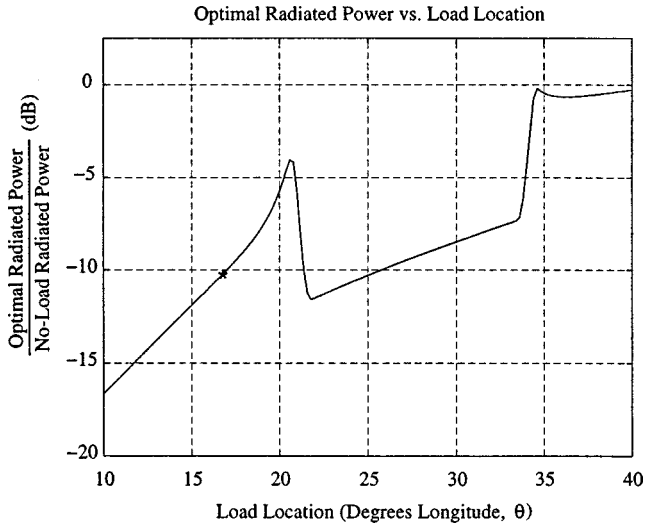


FIG. 6. Numerical experiment 2: dependence of the optimal radiated power—normalized by the no-load radiated power—on load position.

This decrease in reduction is interrupted at $\theta\approx21°$ and $\theta\approx34°$. The upper-left panel of Fig. 3 shows that these values of θ are nulls of the uncontrolled velocity. When the control load is in an angular region on the shell that is vibrating in phase with the disturbance—in this plot, $21°\leqslant\theta\leqslant34°$—the reaction force from the passive load is out of phase with the disturbance, and the patterns easily cancel. When the control load is in a region on the shell that is vibrating out of phase with the disturbance, the reaction force is in phase, and the patterns do not easily cancel.

C. Numerical experiments 3a and 3b: Effect of the passivity constraint on radiation reduction

In experiments 3a and 3b, we modify the passivity constraint to allow the control load to supply power to the structure, which enables examination of the passivity constraint's effect on radiation reduction. We replace the passivity constraint from Sec. II A, $\mathcal{P}_l<0$, with $\mathcal{P}_l<\alpha$. As α increases, we expect increasing reduction in the radiated power until we reach some limit. This limiting value is the radiation reduction for the optimal active load.

The results of experiment 3a are shown in the upper panel of Fig. 7, which plots the unit-disturbance radiated power for (i) the optimal-passive load, (ii) the optimal-active load, and (iii) no-load. In some frequency regions, the active control load reduces radiated power no more than the passive. When the active load does reduce power more than the passive, the difference between the radiation pattern of the load and the radiation pattern of the disturbance allows only a few decibels additional reduction.

The lower panel shows the results of experiment 3b: We plot the optimal radiated power, normalized by the optimal passive radiated power, as a function of α. The lower panel

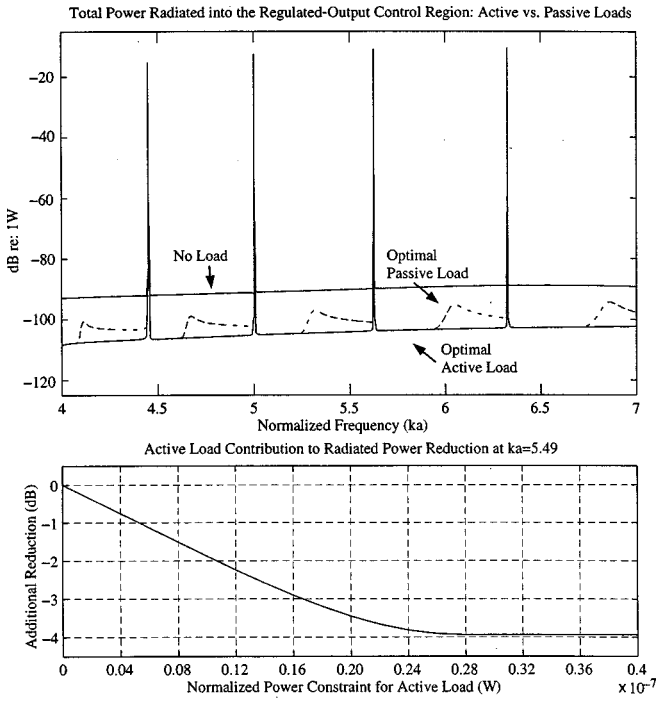


FIG. 7. Numerical experiments 3a and 3b: the optimal-radiated power with and without the passivity constraint, and the optimal tradeoff between radiated power and load power.

shows the optimal tradeoff between radiated power and the power supplied by the active control load. We note that normalized load power greater than $\alpha = 0.28 \times 10^{-7}$ W produces no additional radiation reduction. Similar tradeoff studies are possible for a variety of controller specifications. For example, we could exchange in-sector performance for out-of-sector enhancement.

III. SUMMARY AND CONCLUSIONS

This paper has described a constructive algorithm for bounding the performance of any single-frequency, multiport passive controller. Performance is measured as the maximum singular value of the weighted, closed-loop, disturbance-to-regulated-output transfer matrix, and the algorithm constructs a control load which attains the performance bound.

Using this algorithm, we can design a passive control law for any plant that can be described by a network model, including electrical and electro-mechanical plants. For example, we could apply electro-mechanical coupling devices—like piezoelectric or moving-coil transducers—to the interior of the shell, and we could use the devices' electrical ports as the load points for the network model. In this example, our algorithm would synthesize an electrical load, which in many cases can be implemented more simply than a mechanical load.

Minimizing the maximum singular value of the closed-loop disturbance-to-regulated-output transfer matrix, while maintaining controller passivity, is a constrained optimization problem. We use the network-scattering representation of the load-port variables, and transform the constrained optimization into a semi-definite program. The semi-definite program is solved using a recently available interior-point algorithm.

Our method's utility is illustrated in the design example, control of single-frequency radiated power from a submerged spherical shell. First, we minimize the radiated power over the normalized frequency band $4 \leqslant ka \leqslant 7$: the optimal-passive load reduces radiated power by 10–20 dB, while, in this example, the conjugate-matched load—an often-suggested rule of thumb for passive vibration control—almost always enhances radiated power by a similar amount. Second, we show that the radiated power achieved with the optimal-passive load is sensitive to changes in load position. Third, we remove the passivity constraint to examine its effect on radiation reduction. For the example configuration, in some frequency regions, the optimal-active load often reduces radiated power no more than the optimal-passive load. When the active load does reduce radiated power more than the passive, the additional reduction is only a few decibels.

ACKNOWLEDGMENTS

I am grateful to Professor G. Leonard Tyler and Dr. Alfred Bahr for instructive conversations throughout the preparation of this material. I am also grateful to Professor Stephen Boyd for his assistance with semi-definite programming. This work is supported in part by SRI International.

[1] Y. Nesterov and A. Nemirovskii, *Interior Point Polynomial Algorithms in Convex Programming* (SIAM: Studies in Applied Mathemetics, Philadelphia, PA, 1994).
[2] L. Vandenberghe and S. Boyd, ''Semi-definite programming,'' SIAM (Soc. Ind. Appl. Math.) Rev. **38**, 49–95 (1996).
[3] Ernest S. Kuh and R. A. Rohrer, *Theory of Linear Active Networks* (Holden-Day, San Francisco, CA, 1967).
[4] Leo L. Beranek, *Acoustics* (American Institute of Physics, New York, 1986).
[5] S. Timoshenko and D. H. Young, *Vibration Problems in Engineering* (D. Van Nostrand, New York, 1954).
[6] L. Rogers, ''A damped structure is a friendly structure,'' in *SPIE Vol. 2193, Passive Damping* (SPIE, Bellingham, WA, 1994), pp. 2–5.
[7] J. B. Aldrich, N. W. Hagood, A. von Flowtow, and D. W. Vos, ''Design of passive piezoelectric damping for space structures,'' in *SPIE Vol. 1917, Smart Structures and Intelligent Systems* (SPIE, Bellingham, WA, 1993), pp. 692–705.
[8] R. L. Spangler, Jr. and S. R. Hall, ''Broadband active structural damping using positive real compensation and piezoelectric simultaneous sensing and actuation,'' Smart Mater. Struct. **3**, 448–458 (1994).
[9] N. W. Hagood and A. von Flowtow, ''Damping of structural vibrations with piezoelectric materials and passive electrical networks,'' J. Sound Vib. **146**, 243–268 (1991).
[10] D. G. MacMartin and S. R. Hall, ''Control of uncertain structures using an H^{∞} power flow approach,'' J. Guid. Control Dyn. **14**, 521–530 (1991).
[11] R. J. Benhabib, R. P. Iwens, and R. L. Jackson, ''Stability of large space structure control systems using positivity concepts,'' J. Guid. Control. Dyn. **4**, 487–494 (1981).
[12] R. Lozano-Leal and S. M. Joshi, ''On the design of dissipative lqg-type controllers'' in *Proceedings of the 27th Conference on Decision and Control* (1988), pp. 1645–1646.
[13] M. D. McLaren and G. L. Slater, ''Robust multi-variable control of large space structures using positivity,'' J. Guid. Control Dyn. **10**, 393–400 (1987).
[14] Paul J. Titterton, Jr., ''Network analysis of scattering from an internally loaded spherical shell: Resonance identification and frequency-by-frequency limits on the scattered field,'' J. Acoust. Soc. Am. **98**, 1667–1672 (1995).

[15] Paul J. Titterton, Jr., ''A generalized network formulation for acoustic scattering,'' J. Acoust. Soc. Am. **101**, 107–118 (1997).
[16] Stephen Boyd and L. El Ghaoui, ''Method of centers for minimizing generalized eigenvalues,'' Linear Algebr. Appl. **188–189**, 63–111 (1993).
[17] Lieven Vandenberghe and Stephen Boyd, *SP: Software for Semidefinite Programming: User's Guide* (K. U. Leuven and Stanford University, Stanford, CA, 1994).
[18] Miguel C. Junger and David Feit, *Sound, Structures, and Their Interaction* (The MIT Press, Cambridge, MA, 1986).
[19] R. Herzog, ''Active versus passive vibration absorbers,'' J. Dyn. Syst., Meas., Control **116**, 367–371 (1994).

The effect of sensor placement errors on cylindrical near-field acoustic holography

Gerard P. Carroll
Carderock Division, Naval Surface Warfare Center, 9500 MacArthur Boulevard, West Bethesda, Maryland 20817-5000

(Received 23 January 1998; accepted for publication 10 January 1999)

This paper presents the results of a numerical investigation of the effects of sensor placement errors on the accuracy of cylindrical near-field acoustic holography (NAH). Results are presented using two data sets, one corresponding to numerically generated hologram plane pressures for a periodically ribbed cylindrical shell, and the other to measured hologram plane pressures for an identical small scale (1/50 scale) physical model. The numerically generated hologram plane pressures are valid in the range $ka<4.0$, while the experimental data are valid up to $ka<20$. The motivation for this investigation is to determine whether or not large scale implementation of NAH is feasible given the potentially increased sensor placement inaccuracies which are likely to be encountered at large scale. The effect of both random and systematic axial, circumferential, and radial sensor placement errors are assessed and compared. The majority of the results are presented as wave number comparisons of contaminated and uncontaminated surface pressure reconstructions. It is shown that, as is known to be the case for contamination by background noise, sensor placement errors also result in high wave number noise. Therefore, the success of implementing NAH at large scale when these sensor placement errors are present relies on the selection of a low-pass wave number filter suitable for removing these contaminating factors while preserving the cylinders actual wave number response. © *1999 Acoustical Society of America.*
[S0001-4966(99)04704-9]

PACS numbers: 43.40.At, 43.60.Sx, 43.40.Yq [CBB]

INTRODUCTION

Near-field acoustic holography (NAH), first proposed by Maynard and Williams,[1,2] has been validated in precisely controlled underwater laboratory conditions for small scale planar[3] and cylindrical models. The method provides both surface field and far-field information from near-field pressure measured on a hologram surface. The method has also been demonstrated for irregularly shaped structures (conformal holography)[4] as well as the measurement of energy flow in structures based on reconstructed pressure and velocity.[5,6] Although it is a very powerful structural acoustics tool, large scale implementations of NAH have not been reported. The importance of large scale NAH follows from the fact that small scale models, because of their size, often cannot include the geometric complexities necessary to accurately model the phenomena which occur at large scale. Large scale experiments, on the other hand, must be conducted in large lake or ocean facilities, and therefore, not under the precisely controlled condition for which NAH has been validated. Inexact sensor positioning is the greatest potential source of error in these large scale environments. Background noise, which also may be higher in large scale environments, does not represent a major contaminating source due to the close proximity of the sensors to the radiating structure. This paper reports on the numerical investigation of the effects that sensor placement errors have on the overall accuracy of NAH.

The approach used in this investigation is to simulate a NAH experiment using two different data sets. The first data set consists of analytically generated hologram plane pressures which were calculated using a SARA 2D[7] analytical model of a periodically ribbed cylindrical shell with hemispherical end-caps. The analytical model corresponds to an approximately $\frac{1}{4}$ scale physical model of a submarine. The excitation corresponds to a point force which is located at $0.6L$, where L is the length of the cylinder. The output surface pressure and velocity from the analytical model are used to generate simulated hologram plane data using a numerical implementation of the Helmholtz integral formulation which utilizes a Gaussian quadrature numerical integral formulation.[8] The effect of sensor placement errors is investigated using interpolation methods to calculate hologram plane pressures at locations slightly displaced from the correct measurement location (as determined from the hologram plane measurement grid), and then comparing the holographically reconstructed surface pressure calculated with and without positioning errors. The results of this initial investigation are valid up to *ka* of 4.0. For higher *ka* values, i.e., up to *ka* of 20, a similar study was conducted using actual measured hologram plane pressures for an identical small scale physical model. This experimental data set was obtained at the Naval Research Laboratories (NRL) measurement facility[9] for which the sensor placement errors are known to be small. Therefore, it is possible to investigate the effect of these errors in a manner similar to that used for the simulated hologram plane data set.

I. FORMULATION

A. Near-field acoustic holography

The formulation used in this paper for cylindrical near-field acoustic holography[10] is given in terms of the helical wave spectrum amplitude in Eq. (1):

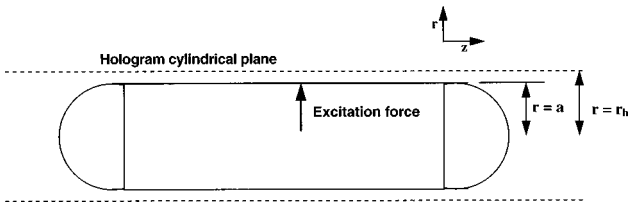


FIG. 1. Diagram showing cylinder, excitation force, and hologram plane.

$$P_n(a,k_z)=\frac{H_n(k_r a)}{H_n(k_r r_h)}P_n(r_h,k_z), \tag{1}$$

where $P_n(a,k_z)$ and $P_n(r_h,k_z)$ are the pressure helical wave spectral amplitudes for the cylindrical surfaces $r=a$ and $r=r_h$ corresponding to the surface of the cylinder under investigation and the hologram plane surface respectively (see Fig. 1). The terms k_z and k_r correspond to the axial and radial components, respectively, of the helical wave spectrum, and are related by

$$k_r=\sqrt{k^2-k_z^2}, \tag{1a}$$

where k is the acoustic wave number. The expression $[H_n(k_r a)]/[H_n(k_r r_h)]$ corresponds to a propagator term which relates the helical wave spectra on the two surfaces to one another. The helical wave spectrum is obtained from the two dimensional Fourier transform of the spatial pressure distribution over the hologram plane, $p(r_h,\phi,z)$, as indicated in Eq. (2):

$$P_n(r_h,k_z)=F_z F_\phi[p(r_h,\phi,z)], \tag{2}$$

where the Fourier transforms are defined by

$$F_z[p(\phi,z)]=\int_{-\infty}^{\infty}p(\phi,z)e^{-ik_z z}\,dz, \tag{3a}$$

$$F_\phi[p(\phi,z)]=\frac{1}{2\pi}\int_0^{2\pi}p(\phi,z)e^{-in\phi}\,d\phi, \tag{3b}$$

and the inverse transforms are defined by

$$F_z^{-1}[P_n(r,k_z)]=\frac{1}{2\pi}\int_{-\infty}^{\infty}P_n(r,k_z)e^{ik_z z}\,dk_z, \tag{4a}$$

$$F_\phi^{-1}[P_n(r,k_z)]=\sum_{-\infty}^{\infty}P_n(r,k_z)e^{in\phi}. \tag{4b}$$

The final result which relates the hologram plane pressure distribution to the reconstructed pressure on the cylindrical test structure is given by

$$p(a,\phi,z)=F_z^{-1}F_\phi^{-1}\left\{\frac{H_n(k_r a)}{H_n(k_r r_h)}F_z F_\phi[p(r_h,\phi,z)]\right\}. \tag{5}$$

A similar formulation for reconstructing the surface velocity $v(a,\phi,z)$ from the hologram plane pressure is given as

$$v(a,\phi,z)=F_z^{-1}F_\phi^{-1}\left\{\frac{-ik_r H_n(k_r a)}{\rho_0 ckH_n'(k_r r_h)}F_z F_\phi[p(r_h,\phi,z)]\right\}, \tag{6}$$

where ρ_0 is the fluid density, c is the speed of sound, and k is the acoustic wave number.

B. Wave number filters

In principal, Eqs. (5) and (6) are sufficient for reconstructing surface pressure and velocity from measured hologram plane pressure. However, any noise in the hologram data tends to be amplified in the backward projection. This is due to the fact that for imaginary arguments (corresponding to subsonic wave mechanisms) the term $[H_n(k_r a)]/[H_n(k_r r_h)]$ increases with wave number, so that at higher wave numbers, where the helical wave number response is typically small, the noise is amplified. The effect of this problem is discussed extensively by Williams *et al.*[3] where it is also shown that the problem is alleviated by low-pass wave number filtering the back-propagated helical wave number response prior to the application of the inverse Fourier transform. Such a filter is intended to eliminate the effects of the noise, while preserving the evanescent response of the cylinder. In this way, the spatial resolution of the reconstruction is not modified by the filtering process. It is further shown by Williams *et al.* that the appropriate value for the wave number filter cutoff frequency cannot be determined *a priori*, but rather, depends on experimental parameters which determine the level of noise contamination, such as the dynamic range of the measurement system, the signal-to-noise ratio, the measurement offset z_h, and the wave number content of the radiator. Typically, low-pass helical wave number filters with rolloff characteristics similar to those given by Eq. (7) are used:

$$\begin{aligned}\Pi(k_r/(2k_c))&=1-\tfrac{1}{2}e^{-(1-|k_r|/k_c)/\alpha}\quad |k_r|<k_c\\&=\tfrac{1}{2}e^{(1-1|k_r|/k_c)/\alpha}\quad |k_r|>k_c,\end{aligned} \tag{7}$$

where k_c is the filter cutoff wave number and α is a parameter which determines the filter roll-off characteristics. The cutoff wave number is given by

$$k_c=\sqrt{\frac{k_{z0}^2+(n_0/a)^2}{2}}, \tag{7a}$$

where k_{z0} and n_0/a are the cutoff wave number components along the helical wave number axes. The case for which $k_z a$ equals an integer n corresponds to n wavelengths around the circumference of the cylinder. Consequently, k_{z0} and n_0/a are numerically equal. Therefore, the cutoff wave number given by Eq. (7a) corresponds to a circular low-pass wave number filter in helical wave number space. Although alternative wave number filtering approaches have been proposed,[11] the one given by Eq. (7) is used by most authors.

A major finding of the present investigation is that, similar to background noise, sensor placement errors also result in high wave number noise. It will be shown that, since all the sensor placement errors considered result in some level of noise in the wave number domain, the success of implementing NAH when these sensor placements errors are present relies on the selection of low-pass wave number filters suitable for removing these contaminating factors while preserving the cylinders actual wave number response. As is the case with background noise, the wave number cutoff fre-

quency cannot be determined *a priori*, but rather, depends on the many contaminating factors which influence the measured hologram plane data.

C. Helmholtz integral equation

The simulated hologram plane data generated from SARA 2D estimates of the test cylinder surface pressure and velocity are obtained using the Helmholtz integral equation as given below

$$\alpha p(\mathbf{r}') = \oiint_{S_0} \left(i\rho_0 ckG(\mathbf{r}:\mathbf{r}')\nu_n(\mathbf{r}) - p(\mathbf{r})\frac{\partial}{\partial n}G(\mathbf{r}:\mathbf{r}') \right) dS_0 . \tag{8}$$

Equation (8) relates the pressure at a field point $p(\mathbf{r}')$ to the pressure and velocity, $p(\mathbf{r})$ and $\nu_n(\mathbf{r})$ on the surface of the cylinder using the free space Greens function $G(\mathbf{r}:\mathbf{r}')$ defined by

$$G(\mathbf{r}:\mathbf{r}') = \frac{e^{ikR}}{4\pi R}, \tag{9}$$

where $R = |\mathbf{r} - \mathbf{r}'|$, i.e., the distance from each source point on the cylinder to each field point on the surface of the hologram plane. Hologram plane pressures were obtained from Eq. (8) using a Gaussian Quadrature formulation[8] to numerically evaluate the integrals.

II. PROCEDURE

A. Numerically simulated data set

SARA estimates of the surface pressure and velocity for 996 axial points and for 16 circumferential orders were utilized in conjunction with Eq. (8) to obtain simulated hologram plane pressures at 256 axial by 128 circumferential locations. Excitation frequencies correspond to $ka = 2.0$ to $ka = 4.0$ in $ka = 0.2$ increments. Using the formulation given in Eq. (5), holographically reconstructed surface pressures were obtained. With no contaminating factors added to the hologram data set, this procedure serves as a validation of the numerical implementations.

B. Effect of sensor placement error

The effect of sensor placement errors was investigated by using interpolation methods to calculate hologram plane pressures at locations slightly displaced from the correct measurement location as determined from the hologram plane measurement grid. Holographically reconstructed surface pressure calculated with and without sensor placement errors were compared to determine their effect. These placement errors are shown in Fig. 2 and correspond to in-plane and radial sensor placement errors. There are many ways in which these placement errors can occur, either individually or in combination, and the ones shown in Fig. 2 are aimed at including most of the important cases.

Figure 2(a), (b), and (c) corresponds to in-plane random placement errors which can occur in the hologram plane. The so-called axial placement random errors shown in Fig. 2(a)

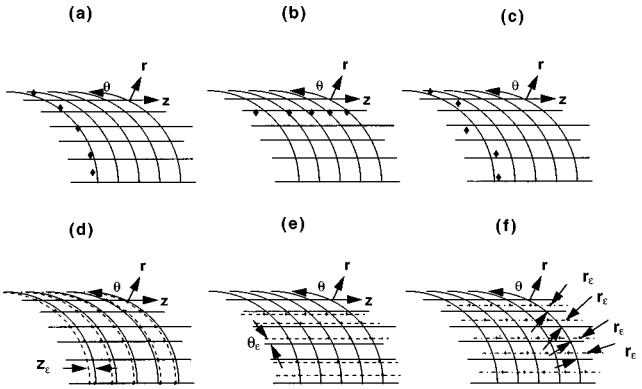


FIG. 2. Sensor placement random errors investigated corresponding to (a) axial placement random error, (b) circumferential placement error, and (c) combined axial and circumferential placement error. Sensor placement bias errors corresponding to a random placement error of line arrays; (d) axial placement bias error z_ϵ, (e) circumferential placement bias error θ_ϵ, and (f) radial placement bias error r_ϵ. Random errors correspond to a normally distributed variation about the true location with a variance of 1 cm.

occur when the sensor position at each location varies randomly along the longitudinal axis, with no variation along the circumferential axis. Similarly, the circumferential placement random errors illustrated in Fig. 2(b) correspond to the sensor position at each location varying randomly along the circumferential axis, with no variation along the longitudinal axis. The combined axial and circumferential placement random errors illustrated in Fig. 2(c) correspond to the sensor position at each location varying randomly along both the longitudinal axis and circumferential axis. In general, this type of error corresponds to errors which might occur if a single sensor is used to acquire the entire hologram plane data set.

The next class of sensor placement errors corresponds to systematic placement errors where an entire axis of sensors is displaced from the correct location. These errors, which will be referred to as bias errors in this paper, actually correspond to random placement errors for the various types of line arrays which might be employed for acquiring the hologram plane data set. For example, Fig. 2(d) shows an axial placement error of a ''hoop array'' of hydrophones. An array such as this could be used to acquire hologram plane data by moving the array along cylinder axis. The so-called axial placement bias error occurs if this array is positioned incorrectly along the longitudinal axis. Similarly, circumferential and radial placement bias errors, shown in Fig. 2(e) and (f), correspond to errors associated with circumferential and radial placement of an axial array of hydrophones. Such an array would collect the necessary hologram plane data by moving the array around the circumference of the cylinder.

Finally, errors associated with an entire cylindrical sensor array are considered. Only one case is considered here. This error which is referred to as a radial offset error corresponds to a systematic eccentricity of the cylindrical sensor array with respect to the model.

To provide a basis of comparison between the various types of placement errors considered, a placement error of 1 cm was considered for all of the data shown. For random errors this corresponds to a normally distributed random error with a variance of 1 cm about the correct placement

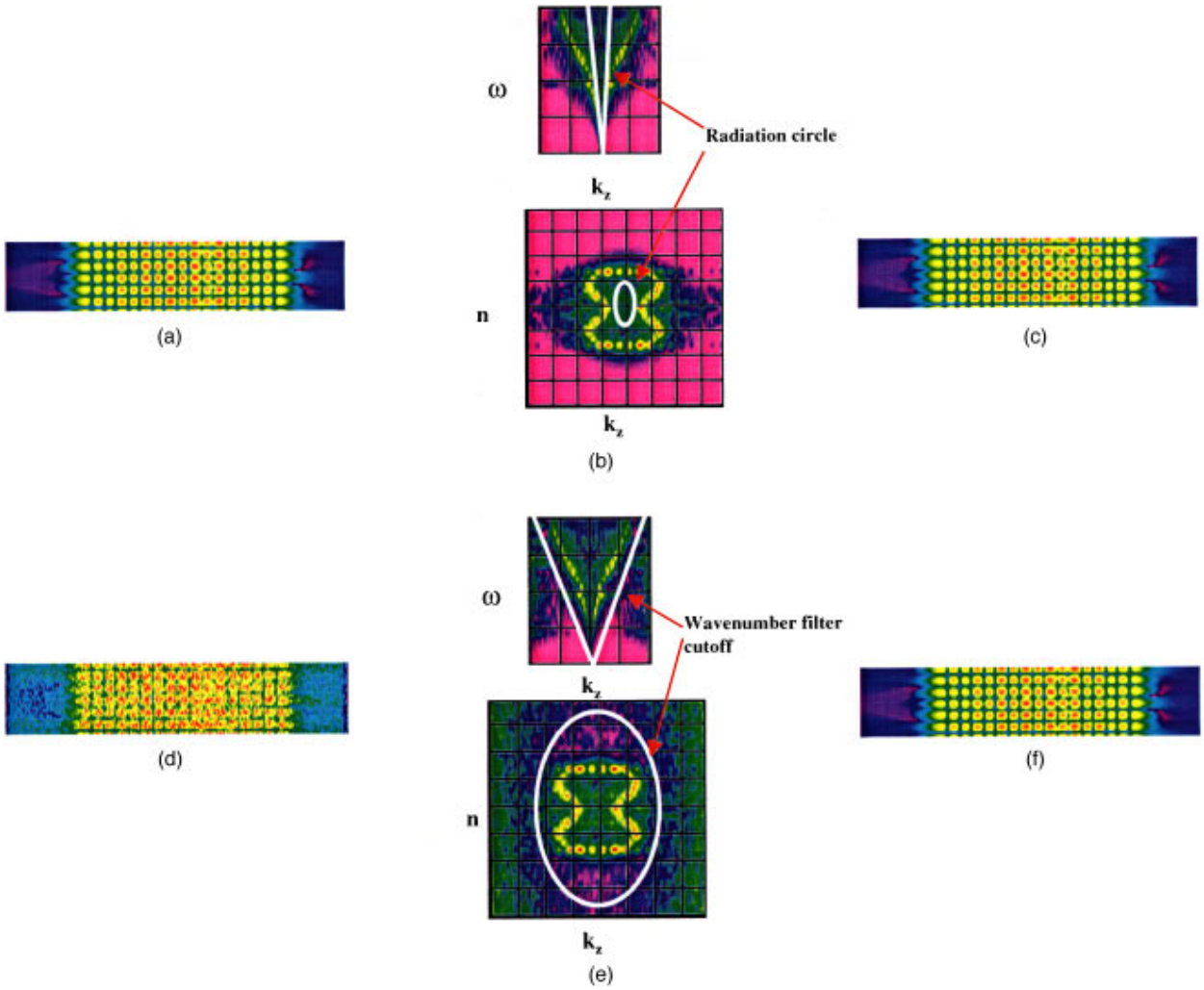


FIG. 3. Graphical representation of the steps involved in the NAH algorithm showing; (a) the uncontaminated hologram plane pressure, (b) the uncontaminated surface pressure reconstructions in the wave number domain, and (c) the reconstructed surface pressure in the spatial domain. The two representations shown in (b) correspond to the frequency versus wave number dispersion curves and the helical wave number response, respectively, and are obtained from the bracketed term in Eq. (5) prior to application of the inverse contaminating influences are added, as shown in (e), (f), and (g). The radiation circle is shown as a reference, and the wave number filter cutoff frequency is shown for the two wave number domain representations used in this study.

location, and for bias errors, to a sensor placement translated 1 cm from the correct placement location. The magnitude of the placement errors considered was based upon an engineering judgment of sensor placement precision and accuracy achievable in large scale facilities, as well as the experience of other researchers doing experiments with large scale models.[12,13]

C. Small scale physical model

Data for the small scale physical model were utilized in a manner similar to the SARA simulated data to determine the effect of various contaminating influences. The obvious assumption in doing this is that these experimental data are relatively free of the contaminating influences under investigation, an assumption which proved to be valid. This physical model is an approximately 1/50 scale model of a submarine. The data set corresponds to 512 axial by 90 circumferential measurement locations. The standoff distance is 6.2 mm (0.244 in.) which is roughly the scale equivalent of a 7.62 cm (3 in standoff) for the $\frac{1}{4}$ scale model. For all the simulations conducted using this data, dimensions were scaled up to correspond to equivalent $\frac{1}{4}$ scale dimensions.

D. Validation of numerical formulations and implementations

Figure 3 shows a graphical representation of the procedure used in this investigation. Figure 3(a) corresponds to the hologram plane pressures. These hologram plane data were obtained alternately from the SARA generated surface data using Eq. (8) and from the small scale physical model data measured at NRL. Figure 3(b) corresponds to the reconstructed surface pressures generated from the hologram plane data using the formulation given by Eq. (5) prior to the application of the inverse Fourier transform. This represents the wave number response of the reconstructed pressure which can be viewed either as frequency versus axial wave number plots or as circumferential versus axial wave number plots as shown. Figure 3(c) shows the reconstructed surface pressure obtained by the application of the Fourier transform. A com-

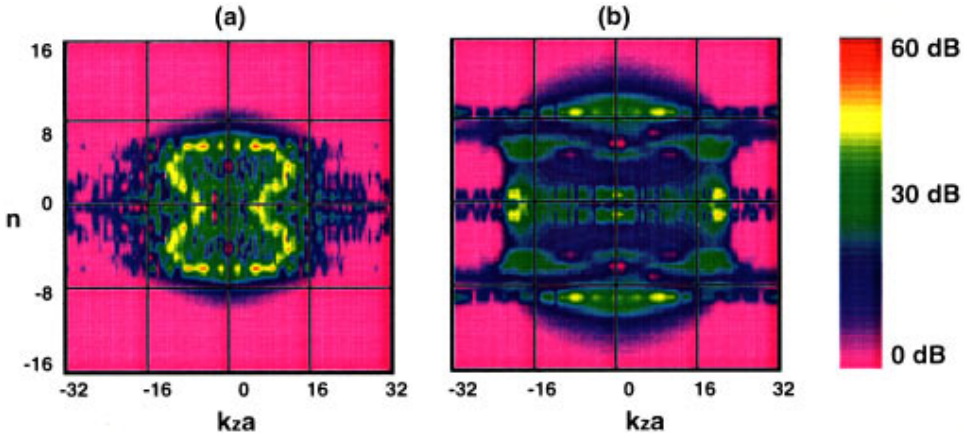


FIG. 4. Helical wave spectra corresponding to uncontaminated holographically reconstructed surface pressure for SARA generated response of ribbed cylindrical model: (a) $ka=2.0$ and (b) $ka=3.6$.

parison of this reconstructed surface response with the SARA generated surface pressures (not shown) serves as a check of the processing algorithms. Comparisons of this type were made and exact correspondence was observed.

Figure 3(d), (e), and (f) corresponds to the results obtained when the same procedure as above is applied to hologram plane data which have been modified by the various contaminating influences under investigation. The effect of these contaminating influences can then be assessed by comparing the uncontaminated and the contaminated reconstructions in either the wave number or the spatial domain. For the reasons given in the next section, these comparisons will be made in the wave number domain in this report.

E. The use of wave number response

For the general example shown in Fig. 3, the effect of the contaminating influence is evident in the comparison of contaminated and uncontaminated wave number response. The radiation circle (which is shown as an elipse due to the scale differences in the k_z and n axes) is shown in the figure as a reference. The effect of data contamination is evident as high wave number ''noise'' which occurs at a sufficiently high wave number so that it can be removed by applying a low-pass wave number filter to the data. Comparing Fig. 3(c) with Fig. 3(f) shows that the spatial domain reconstructed surface pressure for the contaminated data is identical to the

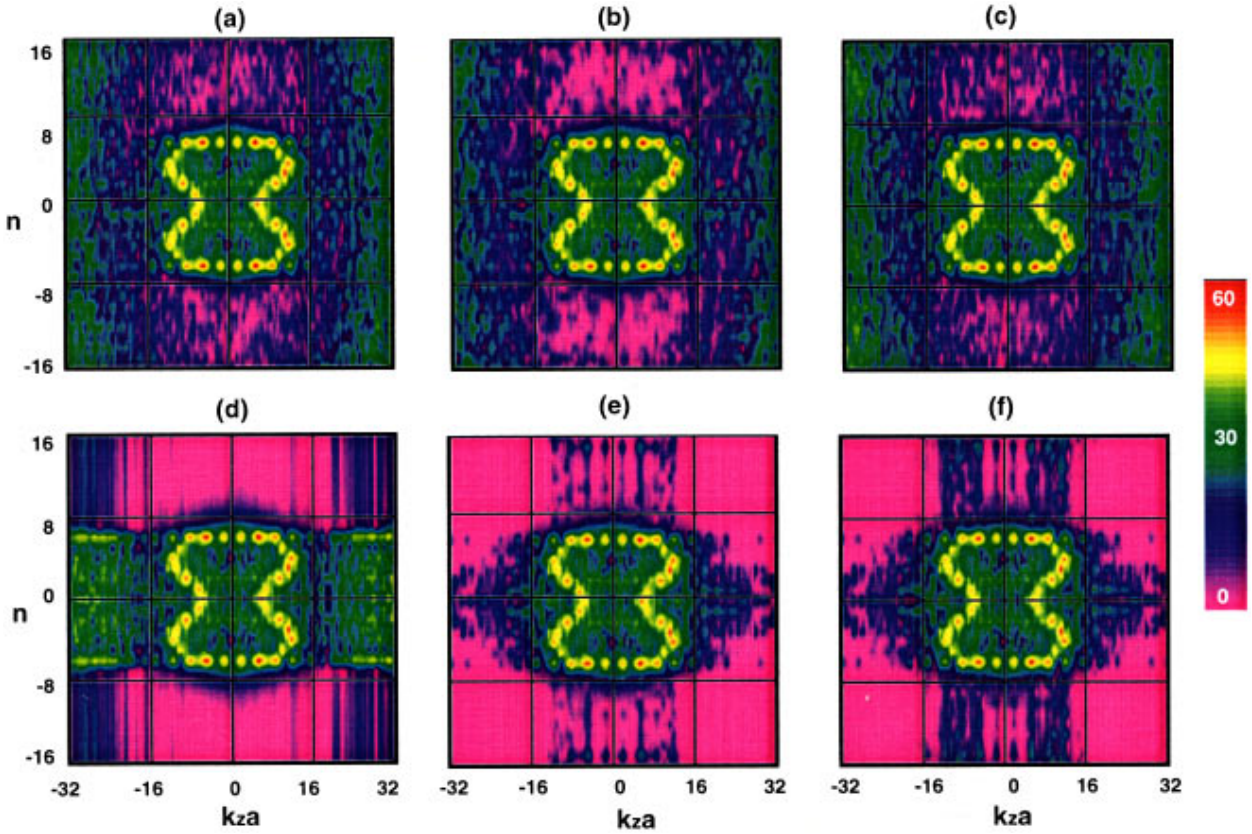


FIG. 5. Same as Fig. 4 for sensor placement errors, $ka=2.0$. Sensor placement errors shown in Fig. 2 correspond to (a) axial placement random error, (b) circumferential placement error, and (c) combined axial and circumferential placement error. Sensor placement bias errors corresponding to a random placement error of line arrays: (d) axial placement bias error, (e) circumferential placement bias error, and (f) radial placement bias error. Random errors correspond to a normally distributed variation about the true location with a variance of 1 cm.

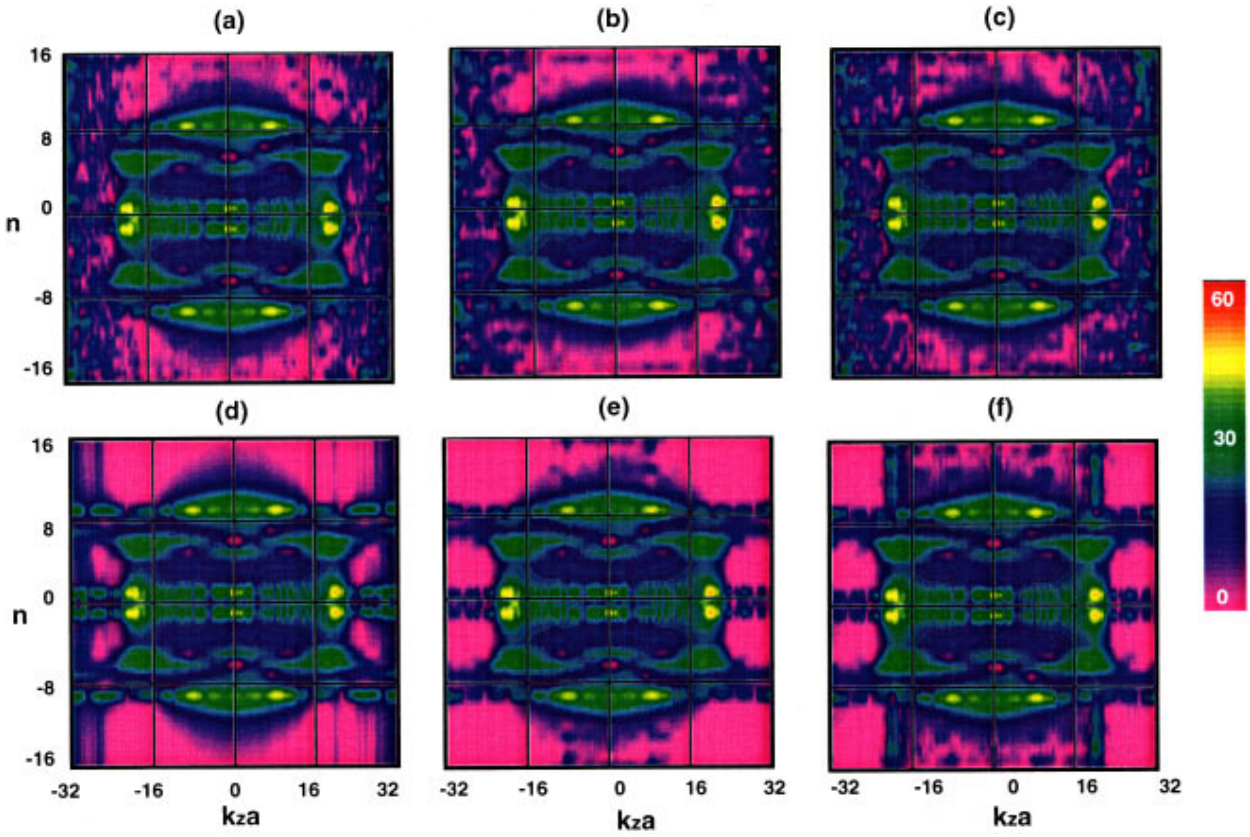


FIG. 6. Same as Fig. 5 for $ka=3.6$.

uncontaminated case after application of the wave number filter. Whereas Fig. 3 represents a general example, it is representative of all the data considered. Therefore, it can be concluded that, if a low-pass wave number filter can be applied to contaminated data so that it eliminates the ''noise'' without significantly modifying the ''signal,'' then that contaminating influence will not have an overall detrimental effect on the holographic reconstructions. In other words, if the two can be separated in wave number space, then holographic reconstructions are feasible. It should be pointed out that in this investigation, the use of simulated data (free from any experimental contaminating influences) makes it possible to determine what is noise and what is signal. This is not always the case in NAH experiments where contaminating influences are present. Usually the practitioner must determine, based on experience and a knowledge of the phenomena participating in the overall response, what part of the data to keep and what part to filter out.

Figure 3 shows that typically the appropriate wave number filter cutoff frequency will increase with frequency. One possible approach for studying the effect of each of the contaminating influences considered is to determine what the appropriate cutoff wave number is for each frequency, and then to use that information to reconstruct spatial surface responses for each frequency. Such an approach requires considerable effort in determining the appropriate filter cutoff wave number, results in comparisons of many reconstructed surface field pressures, and adds a subjective element to the study since the results obtained rely on the investigator's ability to choose the appropriate filter cutoff. A more advantageous approach is to compare the contaminated and uncontaminated wave number response. In this way, using a relatively small number of plots, it is possible to compare the contaminated versus uncontaminated response and determine if it is possible to filter out the noise without greatly modifying the signal. This is the approach which will be employed in the remainder of this paper.

III. RESULTS

A. Hologram plane data set generated from the SARA model

Figure 4 shows uncontaminated helical wave number response (circumferential versus axial wave number) for *ka* of 2.0 and 3.6, respectively. An explanation of some of the structural acoustics mechanisms evident in these figures is given in Ref. 9 and will not be discussed here. For the purposes of this study, the important considerations are whether or not the features shown in the uncontaminated response curves are masked when contaminating influences are added, and whether or not the uncontaminated response can reasonably be separated from the contaminating influences in this frequency/wave number domain by application of a wave number filter.

1. Sensor placement errors

Figures 5 and 6 show the effect of various sensor placement errors on the reconstructed helical wave number response for the selected values of *ka*. A comparison of Figs. 5 and 6 with Fig. 4 indicates that, in general, for all cases

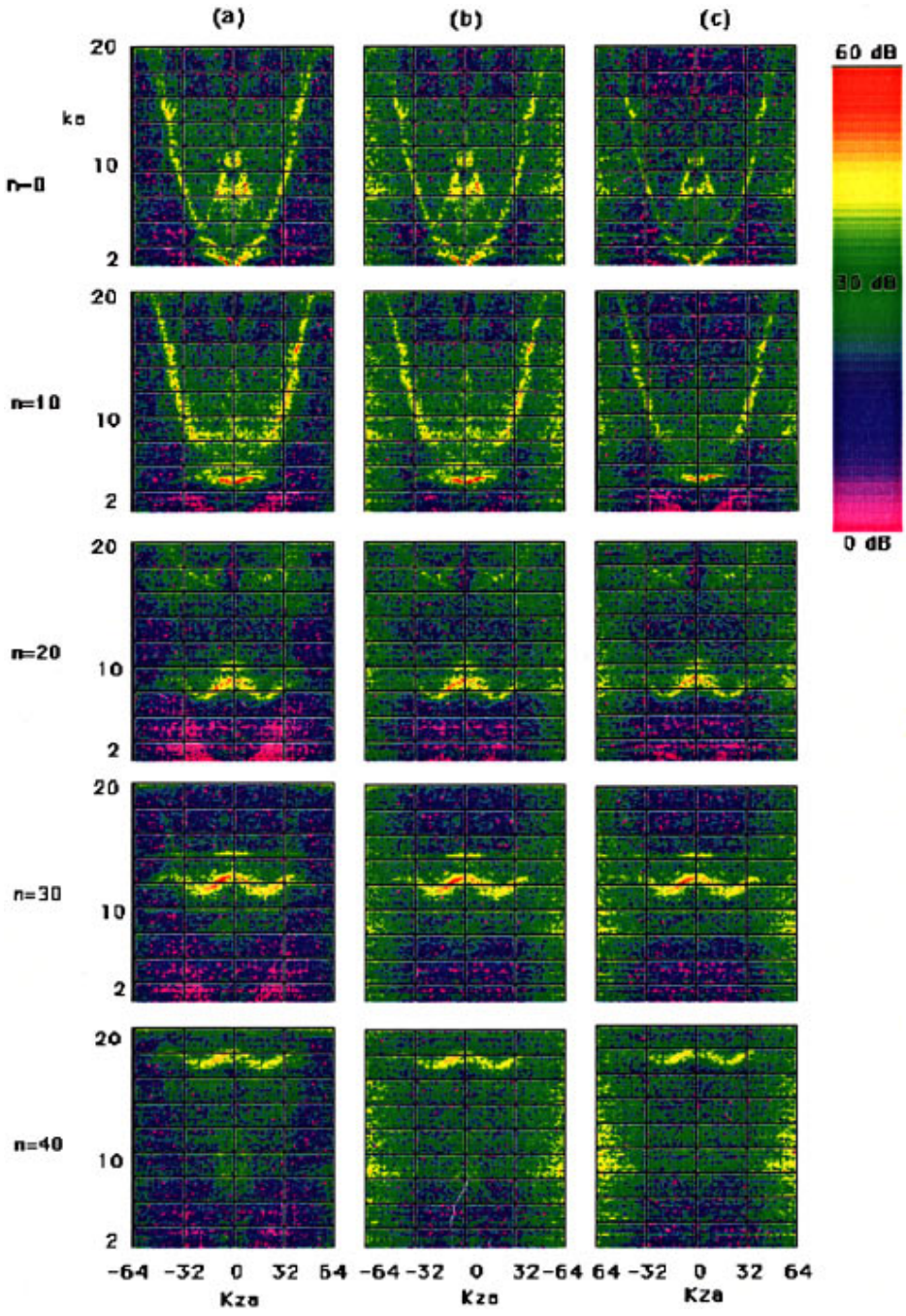


FIG. 7. Reconstructed pressure dispersion curves for the small scale physical model for selected circumferential orders, *n*. Column (a) corresponds to the uncontaminated response, column (b) to the reconstructed response with the addition of a 1-cm axial placement random error [see Fig. 2(a)], and column (c) corresponds to the reconstructed response with the addition of a 1-cm circumferential placement error [see Fig. 2(b)].

shown, the important features of the helical wave number response are preserved when the contaminating factors are added to the data. In general, it appears feasible to filter out the contaminating influences in wave number space without modifying the true response. However, at the higher wave numbers, the wave number separation between the true response and the contaminating factors is often very small. In fact, if one did not know the true response, it may be difficult to determine exactly where to locate the wave number filter cutoff frequency.

This study also provides the possibility of ranking the relative contamination introduced by the various types of sensor placement errors. Such a ranking indicates that the most contamination results from random in-plane errors (axial and circumferential). The results shown in Figs. 5 and 6 indicate that these error mechanisms cause roughly the same level of noise in the helical wave number response. The next highest contamination corresponds to axial placement bias errors which have a much greater effect than circumferential placement bias errors. Circumferential placement bias errors have an effect similar to radial placement bias errors. The smallest effect is due to the so called radial offset errors, which cause no change in the helical wave number response such as those shown in Fig. 4, and is therefore not shown.

B. Small scale physical model results

Figure 7 shows the effect of random axial and circumferential placement errors applied to the small scale physical model data set. Only these two types of errors were consid-

ered based on the results of the previous section, for which it was shown that these two types of errors have the greatest contaminating effect. The results are displayed in terms of frequency versus wave number response at selected circumferential orders. Since this data set corresponds to a maximum *ka* of 20, substantially higher circumferential orders were considered in making these comparisons. Figure 7 shows results for circumferential orders up to $n=40$.

In general, the data lead to the same conclusions as that of the SARA data set. The dispersion characteristics evident in the uncontaminated data are evident in the contaminated response, and the contaminating influences exhibit themselves as high wave number noise. As before, application of a wave number filter to remove the contamination is possible, but may prove to be difficult. For this type of representation, the necessity for frequency dependent wave number filter cutoff frequencies is evident. As was the case with the SARA generated data set, the contamination corresponding to these two types of in-plane random sensor placement errors is about equivalent. The most important conclusion which can be drawn from the studies using this and the previous data set, is that, for errors of the magnitude of those considered in this study, NAH seems to be possible for *ka* as high as $ka=20$.

IV. CONCLUSIONS

The following conclusions can be made on the basis of the simulations described in this paper. It is shown that, similar to background noise, sensor placement errors also result in high wave number noise so that the success of implementing NAH when these sensor placements errors are present relies on the selection of low-pass wave number filters suitable for removing these contaminating factors while preserving the cylinders actual wave number response. As is the case with background noise, the wave number cutoff frequency cannot be determined *a priori*, but rather, depends on the many contaminating factors which influence the measured hologram plane data. In-plane random errors have the largest contaminating influence on holographic reconstructions, followed by in-plane bias errors. Radial random and bias errors have the smallest contaminating influence. Future work aimed at developing improved (and possibly automated) wave number filtering techniques is necessary.

ACKNOWLEDGMENTS

This work was supported by ONR Code 334. The author would especially like to thank Dr. Earl Williams for his assistance with this investigation, and for providing the small scale physical data. The author would also like to acknowledge the contributions of Dr. Geoffrey Main, Dr. David Feit, and Glen Szilagyi who suggested this investigation. Finally, the author would like to acknowledge the support provided by Dr. Karl Washburn and Dr. Ronald Hughes.

[1] E. G. Williams and J. D. Maynard, ''Holographic imaging without the wavelength resolution limit,'' Phys. Rev. Lett. **45**, 554–557 (1980).

[2] J. D. Maynard, E. G. Williams, and Y. Lee, ''Near-field acoustical holography (NAH), I. Theory of generalized holography and the development of NAH,'' J. Acoust. Soc. Am. **78**, 1395–1413 (1985).

[3] E. G. Williams, H. D. Dardy, and R. G. Fink, ''Near-field acoustical holography using an underwater, automated scanner,'' J. Acoust. Soc. Am. **78**, 789–798 (1985).

[4] W. A. Veronesi and J. D. Maynard, ''Digital holographic reconstructions of sources with arbitrary shaped surfaces,'' J. Acoust. Soc. Am. **85**, 588–598 (1989).

[5] E. G. Williams, H. D. Dardy, and R. G. Fink, ''A technique for measurement of structureborne intensity in plates,'' J. Acoust. Soc. Am. **78**, 2061–2068 (1985).

[6] E. G. Williams, B. H. Houston, and J. A. Bucaro, ''Experimental investigation of the wave propagation a point driven, submerged capped cylinder using *K*-space analysis,'' J. Acoust. Soc. Am. **87**, 513–522 (1990).

[7] H. Allik, R. Dees, S. Moore, and D. Pan, ''SARA-2D User's Manual,'' Version 95-3, BBN Acoustic Technologies, New London, CT (1995).

[8] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes, The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1986).

[9] D. Photiadis, B. Houston, and E. G. Williams, ''Wave-number space response of a near-periodically ribbed shell,'' J. Acoust. Soc. Am. **101**, 877–886 (1997).

[10] E. G. Williams, B. H. Houston, and J. A. Bucaro, ''Broadband near-field acoustical holography for vibrating cylinders,'' J. Acoust. Soc. Am. **86**, 674–679 (1989).

[11] S. Hayek and T. Luce, ''Aperature effects in planar nearfield acoustical Imaging,'' ASME J. Vibrational, Acoustics, Stress Reliability Design **110** 91–96 (1988).

[12] A. Clark and P. Watkinson, ''Measurements of under water acoustic intensity in the newfield of a point excited periodically ribbed cylinder,'' *Proceedings of the 2nd International Symposium of Shipboard Acoustics, ISSA'86*, edited by J. Buiten, (Martinus Nijhoff, Dordecht, 1986).

[13] J. A. Clark and D. Feit, ''Time varying wave-number frequency spectra,'' J. Acoust. Soc. Am. Suppl. 1 **84**, S87 (1988).

[14] S. Schreppler and G. Jebsen, ''Intermediate scale measurement system: A state of the art facility for Naval structural acoustics research; system overview and measurement results,'' J. Acoust. Soc. Am. **100**, 2721(A) (1996).

Analysis, testing, and control of a reverberant sound field within the fuselage of a business jet

Robert L. Clark
Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27708

Gary P. Gibbs
Structural Acoustics Branch, NASA Langley Research Center, Hampton, Virginia 23681

(Received 16 May 1996; accepted for publication 24 November 1998)

The objective of this project was to develop a speaker-based active control system capable of dissipating reverberant acoustic energy within the fuselage of a typical business jet aircraft. A ground-based fuselage was utilized as the test-bed for the work conducted in this project. Typical aircraft fuselage trim was installed to more accurately reflect the acoustic properties of a flight-ready aircraft. Speaker locations were selected based upon the acoustic modal characteristics of the fuselage and practical considerations concerning available space in the walls of the trim. Preliminary control experiments involving single-input, single-output, feedback loops resulted in a very localized zone of quiet which yielded no global effects. Multi-input, multi-output feedback control with a distributed array of eight pairs of transducers weighted (spatially) to emphasize energy dissipation of select acoustic modes resulted in as much as 6 dB of global attenuation at acoustic resonances corresponding to the second and third acoustic modes (46 and 71 Hz, respectively) of the enclosure. While objective results were very promising, subjective evaluation of the enclosure resulted in an inaudible difference in the sound field before and after closing the loop on the control system. © *1999 Acoustical Society of America.* [S0001-4966(99)00104-6]

PACS numbers: 43.40.Vn [PJR]

INTRODUCTION

Olson and May (1953) originally introduced the concept of using local feedback control loops for creating local quiet zones and for the control of energy within reverberant sound fields. Realizations of such control systems have proven difficult primarily because of problems associated with high modal density. Clark and Cole (1995) demonstrated in theory that a collocated volume velocity source and pressure transducer can be used to create a dissipative control system. Clark *et al.* (1996) later demonstrated the practical realization of this system with typical audio transducers on an enclosure with low-frequency response dominated by modes in a single dimension. The ideal volume velocity source and pressure transducer can be described as a dual transducer pair which leads to positive real system behavior (passive networks). The frequency response of such dual transducers are characterized by alternating resonances and antiresonances in MacMartin and Hall (1991), and the product of the input and output of dual transducer systems yields units of power. Another example of a dual, collocated transducer pair is the typical point velocity/force transducer paired used in structural control.

To achieve robust control system design and implementation, collocated, dual transducers have been proposed in output-feedback control configurations. Balas (1979) demonstrated the robust potential of direct-rate-feedback control for large space structures with guarantees on stability. MacMartin and Hall (1991) later cast the design such that the power dissipated by the controller can be maximized in an H_∞ sense, leading to a positive real controller with stability guarantees. However, for systems characterized by high modal density, the realization of such controllers proves to yield little performance. The objective typically involves the implementation of a low-authority controller such as a local feedback control loop or multiple local feedback control loops to modify the impedance of the structure. However, as will be discussed in this paper, the closed loop performance of such systems is limited by the relative pole-zero separation, which is negligible in collocated transfer functions for systems with high modal density. Thus, the closed-loop control system is reduced to providing performance locally as opposed to globally.

Both local and global control of an enclosed sound field is presented within this work. In both cases, a single-input, single-output (SISO) control system was implemented in the final design. However, when global control was achieved, the SISO control system was realized through a ''weighted array'' of distributed transducers, wired to act collectively as one largely distributed, collocated transducer. This method of implementation provides a means of emphasizing the selective coupling to acoustic modes over a desired bandwidth and a method of achieving global performance on a system characterized by high modal density. The experimental results presented were obtained from the implementation of a speaker-based active control system designed to dissipate reverberant acoustic energy within the fuselage of a typical business jet aircraft. The introduction serves as a brief summary of the physics associated with the specific problem under consideration: *active control of a reverberant sound field.*

I. THE REVERBERANT SOUND FIELD

The propagation of sound within an enclosure is the result of sound radiation from the source and sound which is reflected or scattered by the walls and objects which constitute the enclosure. When acoustic waves are reflected, the field is described as reverberant. If one could control the aircraft fuselage and constrain the enclosure to emulate an anechoic room, the reverberant response to stochastic external noise sources such as turbulence or other broadband spectra would be negligible. At mid to high frequencies, significant levels of sound absorption occur in the fuselage due to the trim, carpet, seats, and passengers. However, at low frequencies (below 250 Hz), very little absorption occurs and the acoustic modes of the fuselage tend to ''ring.'' This ringing serves to amplify stochastic noise sources such as broadband engine noise or turbulent boundary layer noise. The emphasis in this section is to review two fundamental theories of reverberant sound fields and describe the boundaries between the two theories for physical application and interpretation of the reverberant sound field within the fuselage of a business jet aircraft.

A. Sabine–Franklin–Jaeger theory

Sabine discovered that the sound within a reverberant room can be described in terms of an average energy density per unit volume and is approximately uniform throughout the room. The theory is based upon reverberant enclosures with a slow rate of sound energy dissipation. In other words, the time required for the sound waves to propagate between the walls of the enclosure is much less than the time required to dissipate a significant portion of the energy (a good assumption in the fuselage at low frequencies). The conservation of energy within an enclosed sound field can be stated mathematically as follows (Pierce, 1981):

$$\frac{d}{dt}\int_V \underbrace{\left(\frac{1}{2}\rho_0 v^2+\frac{1}{2}\frac{p^2}{\rho_0 c^2}\right)}_{\text{energy density}} dV = \underbrace{P_{\text{in}}-P_{\text{diss}}}_{\text{nonconservative terms}},$$

where the time rate of change of the energy density over the volume of the enclosure is equal to the sound power supplied by all sources less that dissipated. Upon assuming that the average energy density is uniform over the volume and describing the average power dissipation in terms of an equivalent area of ''open windows,'' a first order-differential equation results which can be solved in terms of a single time constant.

Another important parameter obtained from the analysis of reverberant sound fields is the *radius of reverberation*. The radius of reverberation is the radius at which the sound pressure level due to the direct field of the source is equal to that due to the reverberant field. Very near the source (as is typical in active control applications), the assumptions made in the Sabine–Franklin–Jaeger theory are invalid. The radius of reverberation serves to define a minimum separation distance between noise source and microphones used in reverberant field measurements.

For collocated control of enclosed sound fields, each transducer pair (microphone and speaker) is always separated by a distance much less than the radius of reverberation, which implies that the sound measured at the microphone is dominated by the direct field of the control source. For systems displaying high modal density at low frequency such as the enclosed sound field, using multiple single-input, single-output collocated control sources within the reverberant sound field will result in local as opposed to global control as will be demonstrated in Sec. III.

B. Modal theory

A detailed mathematical model of the reverberant sound field can be developed by solving the homogeneous wave equation for a room subject to rigid wall boundary conditions in terms of a series expansion which results in an eigenvalue problem. The difficulty with this type of solution is that the modal density of the enclosure dictates a series expansion which must include a very large number of terms to accurately capture the dynamics of the system, even at low frequencies where there is significant modal separation.

The need for including higher-order modes in a low-frequency model is not immediately apparent and is typically ignored in a number of analytical studies implementing control systems. Typical implementation would include the first few modes outside the bandwidth of interest which is inadequate to capture the dynamics present in an actual system for collocated systems. This paper demonstrates the need to include significantly more terms in the modal expansion for collocated cases. The higher-order modes affect the low-frequency response of a reverberant sound field most when the transducers are collocated because, in the limit, all of the modes are in-phase at the point of collocation and thus contribute in an additive sense to the acoustic response at that field point. If the transducers are separated by a distance exceeding the radius of reverberation, the relative phasing and amplitude contribution of higher-order modes becomes random in nature and thus significant cancellation between pressures induced by higher-order modes occurs. Thus, for noncollocated transducers, fewer modes are required in the series expansion to accurately capture the frequency response at low frequencies.

As detailed in prior work by Bishplinghoff and Ashley (1975), the Laplace transform of the response of a reverberant system can be expressed in terms of a series expansion of a finite number of modes and that of all higher-order modes:

$$y(s)=\sum_{n=1}^{N}\frac{Q_n(s)\Gamma_n}{s^2+\omega_n^2}+\sum_{n=N+1}^{\infty}\left(\frac{Q_n(s)\Gamma_n}{\omega_n^2}-\frac{s^2 q_n(s)\Gamma_n}{\omega_n^2}\right),$$

where $y(s)$ is the Laplace transform of the output which can represent pressure in a reverberant sound field, $Q_n(s)$ is the Laplace transform of the generalized force, Γ_n is the modal participation corresponding to the spatial location of the output transducer, $q_n(s)$ is the response in generalized coordinates, and ω_n is the natural frequency of the nth mode. The first term *in the second series expansion* above represents the low-frequency displacement contribution of the out-of-bandwidth modes, and the second term represents the inertial

contribution of the out-of-bandwidth modes. To obtain the frequency response, it is typical to let $s=j\omega$. In so doing, one observes that the *second term in the second series expansion is negligible* as long as $\omega \ll \omega_n$. For collocated transducers, $Q_n(s)=\Gamma_n u(s)$, and the transfer function can be expressed as follows:

$$\frac{y(s)}{u(s)}=\sum_{n=1}^{N}\frac{\Gamma_n^2}{s^2+\omega_n^2}+\sum_{n=N+1}^{\infty}\frac{\Gamma_n^2}{\omega_n^2}=\hat{G}(s)+K_{N+1}^{\infty}.$$

As indicated, the frequency response can be expressed in terms of a series expansion of a finite number of modes and an infinite series expansion of the displacement contribution of out-of-bandwidth modes. The first series is represented by an approximation of the plant transfer function, $\hat{G}(s)$, while the second series expansion can be represented by a single constant, K_{N+1}^{∞} as discussed by Clark (1995). In terms of physical understanding, the low-frequency, displacement contribution of out-of-bandwidth modes is essential to predicting what is typically referred to as near-field dynamics in wave domain models. The ''near-field'' dynamics are an essential element in the prediction of closed-loop system performance because these dynamics dictate the positions of the system transfer function zeros and thus the ultimate destination of the system poles.

The assumptions made in computing the reverberation time of an enclosed sound field (based on Sabine–Franklin–Jaeger theory) are not applicable to the fuselage of a business jet at low frequencies. Thus, a modal approach can be formulated which relates the decay of sound energy to the equivalent damping (Kuttruff, 1979). Since the response of each acoustic mode of the enclosure can be expressed in terms of the time constant associated with the natural frequency and damping ratio, the reverberation time can be computed as follows for a particular mode:

$$T_{60}=\frac{3\ln(10)}{\xi_n\omega_n},$$

where ξ_n is the damping ratio of the nth acoustic mode, and ω_n is the natural frequency of the nth acoustic mode. This expression was obtained from an average of the energy density which neglects the cross-coupling of modes when the damping ratio is low (less than 10%).

It is obvious that certain assumptions apply to the Sabine–Franklin–Jaeger theory of room acoustics and limitations apply to the applicability of a modal model to enclosures with high modal density (almost all enclosures except ducts at low frequency). Regardless, the transition in applicable modeling is not discrete from one approach to the other, but rather continuous and knowledge of each modeling approach and assumptions made provides physical insight into the reverberant response of enclosed sound fields. The fuselage of a business jet aircraft is an enclosure which lends itself to modal modeling at low frequencies and statistical modeling at high frequencies.

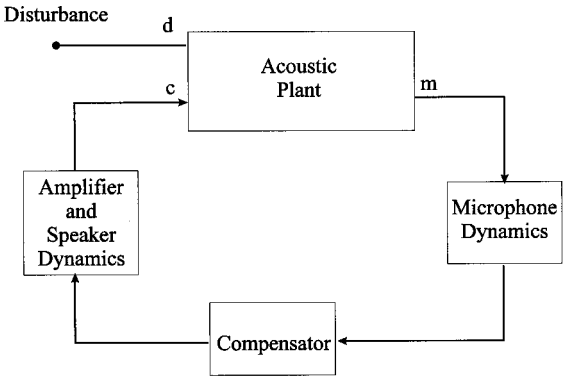


FIG. 1. General block diagram of control system.

C. Implications on the control of enclosed sound fields

Control system design is often formulated in terms of a system model. However, for the enclosed sound field, the modal density is such that a complete dynamic model is impractical. A low-frequency modal model can be constructed, including the ''displacement'' contribution of out-of-bandwidth modes, and this leads to an accurate prediction of the system zeros which is required to predict the performance of the closed loop system. The transfer function resulting from collocated transducers in a reverberant acoustic fields produces alternating poles and zeros at low frequencies, which is desirable for robust stability; however, performance is limited because all modes are in phase for collocated transducers and thus any control energy expended on the control of the system is distributed among an infinite number of modes. Hence, very little performance can be obtained at any single mode. The solution to this problem rests in the development of distributed collocated transducer arrays and will be outlined in the control section.

II. THE FUSELAGE ENCLOSURE: ANALYSIS AND EXPERIMENT

In the following subsections, the fuselage enclosure is reviewed with a preliminary analysis and a discussion of the experimental results. The limiting physics for local and global control of sound fields are presented in the following subsections.

A. A description of the general system

A block diagram of a general acoustic system is presented in Fig. 1. For the aircraft enclosure used in this study, a series of speakers was mounted in the trim of the fuselage. The speakers were accompanied by a substantially collocated microphone at each speaker. A single speaker/microphone pair was selected for the single-input, single output analysis and a full array was used for the multi-input, multi-output (MIMO) analysis. The outputs of the microphones are passed through a compensator, dynamic or static, and subsequently through power amplifiers to drive each speaker. The system was subjected to an unwanted disturbance at a location far from the collocated control transducer pair which is represented schematically in Fig. 1.

TABLE I. Morel speaker properties.

Variable	Value	Units
Resonant frequency	32	Hz
Bl	5.11	N/A
L	0.47	mH
Rdc	3.6	ohms
D	8	in.
Mms	28	gm
Qms	1.56	
Qes	0.79	

The microphones used in this study were custom designed from Panasonic electret condenser elements, and the corresponding power supplies were developed to provide a high sensitivity and operate on aircraft power systems, 28 V dc. The microphone and power supply system can be modeled with three poles and one zero (Clark *et al.*, 1996).

The speakers used in this study were manufactured by Morel of England, model number MW 224. These drivers are 8 in. in diameter and were selected because they have an extremely shallow mounting depth requirement of 2 in. with a bandwidth extending from 28 to 3000 Hz. These speakers can be flush mounted within the trim and fit within a frame bay of a typical business jet aircraft while providing the desired frequency response characteristics. The properties of the speakers are listed in Table I.

B. Physical insight from an analytical model

In this section a simple model of a three-dimensional sound field will be presented. For the purposes of this model a simple rectangular room (rectangular parallelepiped) will be used with dimensions approximately equal to that of the interior of the business jet fuselage used in this study. This cavity was chosen because of its simplicity, and it demonstrates the relevant physics. A diagram of this model is presented in Fig. 2. The eigenfunctions of this ideal enclosure are of the form:

$$\varphi(\mathbf{x},l,m,n)=\Phi_{lmn}\cos\left(\frac{l\pi x}{L_x}\right)\cos\left(\frac{m\pi y}{L_y}\right)\cos\left(\frac{n\pi z}{L_z}\right),$$

where Φ_{lmn} is the modal amplitude, l is the modal index for the x direction, m is the modal index for the y direction, n is the modal index for the z direction, L_x is the length of the room in the x direction, L_y is the length of the room in the y direction, and L_z is the length of the room in the z direction. The root of the eigenvalues (natural frequencies) are given by the following relation:

$$\omega_{lmn}=c\sqrt{\left[\left(\frac{l\pi}{L_x}\right)^2+\left(\frac{m\pi}{L_y}\right)^2+\left(\frac{n\pi}{L_z}\right)^2\right]},$$

where c is the speed of sound in air.

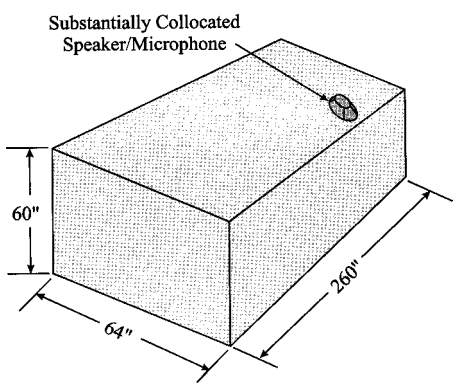


FIG. 2. Schematic of rectangular cavity model.

TABLE II. Analytical resonant frequencies.

Mode No.	Resonant frequency (Hz)
1	0
2	26
3	52
4	77
5	103
6	105
7	108
8	112
9	115
10	117
11	123
12	129
13	130
14	136
15	147

For this particular model, the room had dimensions of $64\times60\times260$ in. $3h$ and the speaker was mounted as shown in Fig. 2. The resulting resonant frequencies of the first 15 modes of this system are presented in Table II. Note the rapid increase in modal density above 100 Hz.

The frequency response of the collocated transducer pair for the system depicted in Fig. 2 is presented in Fig. 3 for a 10-mode model, a 50-mode model, and a 50-mode model including the compliance of 64 000 out-of-bandwidth modes, respectively. Note the subtle increase in the magnitude of the frequency response when the compliance of the 64 000 out-of-bandwidth modes is added to the 50-mode model of the collocated transducer pair. As discussed earlier and as seen in Fig. 3, the most significant effect of the out-of-bandwidth modes at low frequency is observed in the migration of the zeros (antiresonances) of the frequency response. For the

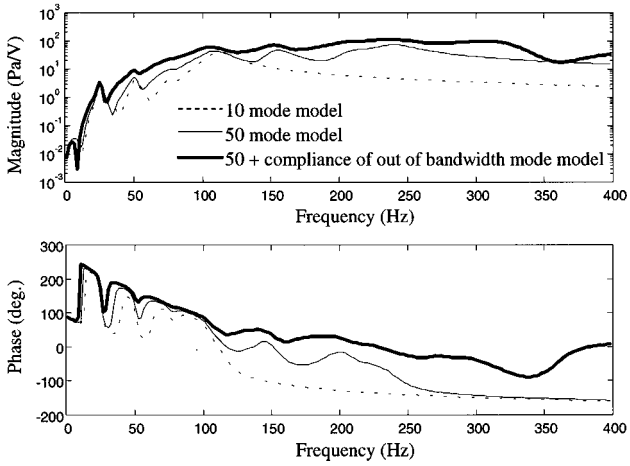


FIG. 3. Collocated frequency response of rectangular enclosure.

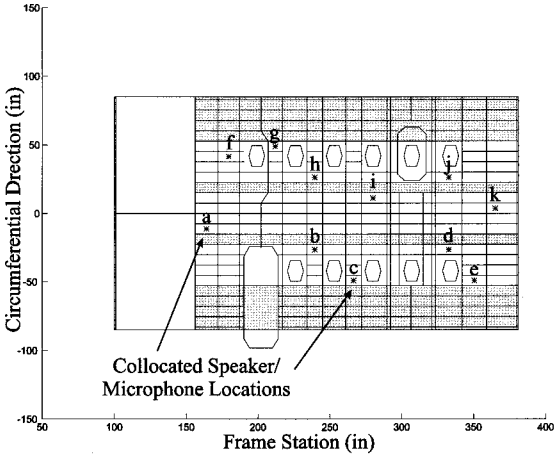


FIG. 4. Speaker and microphone locations.

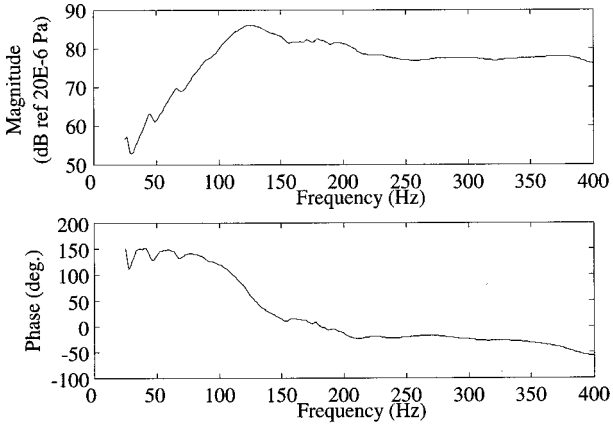


FIG. 5. The SISO loop transfer function, speaker *d* to microphone *d*.

modal model which includes the compliance of out-of-bandwidth modes, the poles (resonances) and zeros (antiresonances) are very close.

The fundamental reason for this observation is that collocated transducers have a significant near-field component which requires an excessive number of modes to accurately capture the system response. In modal terms, since the modes of a collocated transducer pair are all in-phase at the point of collocation, one must expand the series for a very large number of modes to reach convergence of a system with high modal density. However, for the non-collocated transducers, as the separation distance approaches or exceeds the radius of reverberation, the relative phase between modes becomes random in a statistical sense and only the ''long-wavelength,'' low-order modes have similar phase. Thus, the reverberant field response (due to reflections) is comparable or greater than the direct field response (due to the source) and hence fewer modes are required to accurately capture the system frequency response at low frequency.

C. Experimental results

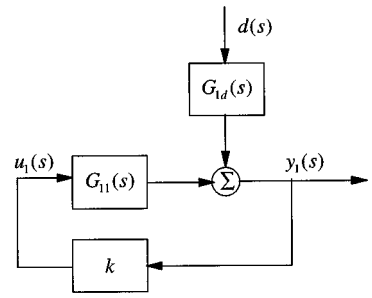
In this section, the results of the experimental transfer function measurements are presented and compared when applicable to the previously discussed model. The aircraft fuselage used in this experiment has a cabin 72 in. in diameter and length between pressure bulkheads of 260 in. The fuselage has frames and stringers as shown by black lines in Fig. 4. The aircraft was outfitted with ten Morel speakers mounted in the frame bays as shown in Fig. 4 and represented by the letters a–k. The fuselage has been unwrapped such that 0 on the vertical axis represents the top center seam of the aircraft, and note that the windows, door, and escape hatch are also outlined in Fig. 4. Also ten microphones where mounted 1 in. from the speakers shown in Fig. 4. The speakers were excited with band-limited white noise from 0 to 400 Hz and the resulting transfer function between speaker and collocated microphone pair *d* is shown in Fig. 5. This transfer function is typical of all collocated measurements.

As is evident from Fig. 5, the system displays relatively high damping (5%–7%) even at the low frequencies, and the zeros are very near the poles for this case (typical of all drive point cases). A combination of the high modal density and high damping causes the modes to be indiscernible above 100 Hz. The general shape of the transfer function should be compared with Fig. 3 (dark line). Modal models as presented in Sec. II B must include the stiffness contributions for the out-of-band modes to accurately capture the true transfer function. Use of low-order modal models to predict closed loop control performance in this case (collocated) will significantly exaggerate performance achievable in a test.

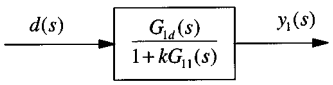
III. CONTROL APPROACHES AND RESULTS

There are two basic control approaches which can be employed in the control of enclosed sound fields: local control and global control. Nelson and Elliott (1992) described the local control approach in terms of ''zones of quiet,'' whereby the objective of the control system is to reduce the sound pressure level in the vicinity of the error microphones. The radius of influence is limited by the shortest wavelength of sound which must be controlled, and considerable spillover can result from this control approach. However, zones of quiet can be achieved with feedback control as outlined originally by Olson and May (1953). In more recent years, Nelson and Elliott (1992) have discussed the local control approach with respect to adaptive feedforward control.

The alternative to local control is global control. In global control of enclosed sound fields one seeks to modify the dynamics of the enclosure, mathematically, the eigenvalues of the closed-loop system. The primary method for accomplishing this task is through the modification of the system input impedance to exogenous inputs. Through feedback, both the stiffness and damping associated with the closed-loop system can be modified. However, the closed-loop system performance is limited by the relative location of the open-loop zeros and poles of the loop transfer function. For reverberant systems with high modal density such as an acoustic enclosure, the zeros of any single collocated loop transfer function are nearly coincident with the poles as demonstrated in Sec. II. Thus, the performance of the closed-loop control system is severely limited. Spatial signal processing can be utilized in conjunction with substantially (with respect to a wavelength) collocated transducers to increase the



Block diagram of SISO system



Simplified block diagram of SISO system

FIG. 6. Block diagram of SISO implementation.

separation between poles and zeros for improved closed-loop performance. Each control approach and experimental results from tests employed on a typical business jet fuselage are documented in the following sections.

A. Single-input, single-output feedback control

Consider the block diagram of the SISO system presented in Fig. 6. The purpose of this block diagram is to illustrate the path associated with the exogenous disturbance and why local control results for high modal density systems. As illustrated, a disturbance, $d(s)$, serves to drive the reverberant plant through a transfer function $G_{1d}(s)$ relating the response at error sensor 1 to the disturbance input. For linear systems, the total response at error sensor 1 can be obtained from the sum of that due to the disturbance and that due to the control. The transfer function $G_{11}(s)$ relates the control input $u_1(s)$ to the output at error sensor 1, and, as illustrated, the control signal is obtained by direct output feedback control in this case with a scalar gain, k. For the purpose of this simple example, transduction device dynamics are not considered. The loop transfer function is assumed to be obtained from collocated, dual transducers such that an infinite gain margin exists. Since the objective is to attenuate the disturbance, the block diagram presented in Fig. 6 is reduced to a simplified block diagram relating the system output at error sensor 1 to the disturbance.

As illustrated, the system transfer function can be expressed as follows:

$$\frac{y_1(s)}{d(s)}=\frac{G_{1d}(s)}{1+kG_{11}(s)}.$$

If the loop transfer function, $kG_{11}(s)$, is large over the bandwidth of the disturbance, then the system transfer function can be approximated as follows:

$$\frac{y_1(s)}{d(s)}=\frac{G_{1d}(s)}{kG_{11}(s)}.$$

Since the poles are common to the disturbance and control system transfer functions, the closed loop transfer function can be expressed as a function of the gain and the numerator polynomial associated with the disturbance and control transfer functions, $N_{1d}(s)$ and $N_{11}(s)$, respectively:

$$\frac{y_1(s)}{d(s)}=\frac{1}{k}\frac{N_{1d}(s)}{N_{11}(s)}.$$

Thus, as the gain is increased, the response at the error sensor decreases in the limit of high gain; however, in terms of global control, notice that the closed-loop poles have now been assigned to the zeros of the loop transfer function. *For reverberant plants with high modal density, the zeros of collocated transfer functions take on nearly the same values as the poles of the transfer function and thus the global effect of this control system is negligible*. To exercise global control, the loop transfer function must be modified such that the closed-loop poles can be move further into the left-half s-plane (increased stiffness and damping).

Consider the following expression for the poles of the loop transfer function:

$$D(s)=\prod_{i=1}^{\infty}(s^2+\omega_i^2),$$

where the complex conjugate poles, $\pm j\omega_i$, are undamped. For the acoustic enclosure, there are an infinite number of poles and thus the denominator of the transfer function can be expressed in terms of the product expansion. The objective is to increase the energy dissipation of the system over a desired bandwidth. To accomplish this task, the poles of the system must be modified. Let us assume that the loop transfer function can be modified such that

$$G_{11}(s)=\frac{s}{s^2+\omega_n^2}.$$

The response of the loop transfer function is thus limited to that of the nth mode of the reverberant system, and for the purpose of this simple example, the dissipative term is ignored. From the block diagram presented in Fig. 6, the equivalent system transfer function from the disturbance to output can now be expressed as follows:

$$\frac{y_1(s)}{d(s)}=\frac{G_{1d}(s)}{1+ks/(s^2+\omega_n^2)}$$

$$=N_{1d}(s)/(s^2+ks+\omega_n^2)\prod_{i=l,i\neq n}^{\infty}(s^2+\omega_i^2).$$

As indicated, for this loop transfer function, the control gain serves to enhance the energy dissipation associated with the nth mode of the system. In the limiting case, ''modal'' sensors and actuators would be required to achieve this ideal control scenario. However, arrays of transducers can be distributed to increase the residue associated with desired poles and thus place greater emphasis on control system performance with respect to desired modes over the intended bandwidth of operation. Multi-variable control approaches are required to accomplish this task.

B. Multi-input, multi-output feedback control

For the purpose of this work, a square control system will be considered (equal number of inputs and outputs), and the transducers are substantially collocated (with respect to the acoustic wavelength over the bandwidth of interest) for

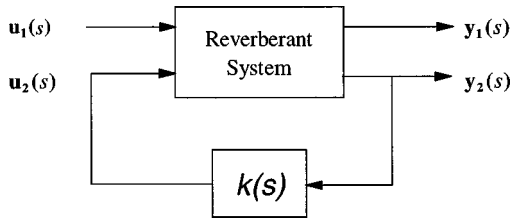


FIG. 7. Block diagram of MIMO system.

enhancing stability robustness. The emphasis is placed upon the limiting physics associated with the high modal density, reverberant system in MIMO control.

The reverberant system can be expressed in terms of the standard two-port problem as illustrated in Fig. 7. The inputs and outputs are vectors and display the typical nature of a MIMO system. The exogenous inputs are represented by $\mathbf{u}_1(s)$ and the control inputs are represented by $\mathbf{u}_2(s)$. The outputs to be minimized are contained in the vector $\mathbf{y}_1(s)$, and the outputs available for feedback control are contained in the vector $\mathbf{y}_2(s)$. The objective is to determine the optimal compensator $\mathbf{k}(s)$ such that an appropriate norm of the loop transfer function between $\mathbf{y}_1(s)$ and $\mathbf{u}_1(s)$ is minimized. In practical terms, one seeks to determine a compensator which modifies the global dynamics of the system over the bandwidth of interest without requiring excessive levels of control effort.

If independent modal space control (Meirovitch and Baruh, 1980) were an option, then one could arbitrarily assign the impedance of each mode of the reverberant system as discussed in Sec. III A, and the control problem would be solved. This approach suffers from a practical method of synthesizing complex modes of real systems. However, if an array of substantially collocated transducers are employed, then the outputs and inputs of the transducers can be weighted so as to emphasize the control of select reverberant modes in the bandwidth of interest. The objective is not necessarily to create ''modal filters,'' but rather to simply weight the input and output arrays such that the zeros are well spaced with respect to the poles of modes to be controlled. Separating the poles and zeros leads to enhanced closed-loop performance and global control. The stability characteristics associated with collocated transducers can be preserved if the inputs and outputs of the dual, collocated transducers are weighted in an identical fashion, reducing the multi-variable transducer array to a distributed, single-input, single-output pair of collocated, dual transducers.

C. Application of active control to the fuselage

The SISO and MIMO control approaches discussed previously were implemented using analog and digital hardware on a ground-based, business jet fuselage. In all control experiments, a dynamic compensator was implemented to modify the open-loop frequency response obtained between the microphones and speakers. The dynamics of the speaker and microphone significantly limit the bandwidth for output feedback control, and, as such, phase compensation is required. The use of phase compensation to increase the bandwidth and reduce spillover for output feedback control of

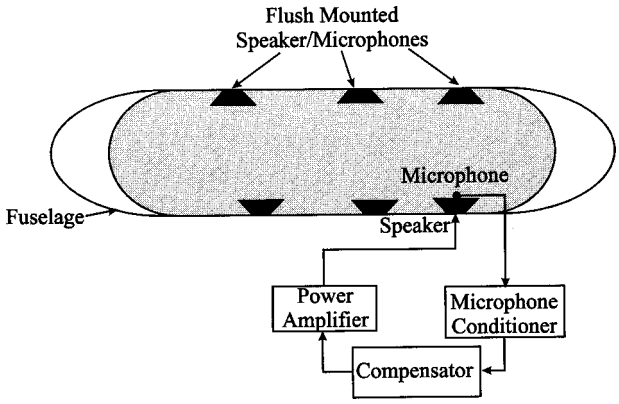


FIG. 8. Schematic diagram of fuselage and instrumentation.

enclosed sound fields was previously detailed by Clark *et al.* (1996) and in the interest of conciseness will not be repeated here. The reader is encouraged to review the reference for details of dynamic compensation as the same dynamic compensator was used for the SISO and MIMO applications discussed here. The primary difference between the two-alternative control system implementations is that the SISO controller utilized a single microphone and speaker while the MIMO controller utilized an array of microphones and speakers, weighted to emphasize control of the low-frequency modes of the enclosure and wired to respond as a single transducer pair. In each case, once the dynamic compensator was implemented to modify the phase of the open loop system, a cost functional was constructed to balance control effort with the reduction in sound pressure at each of the microphones. The optimal gains (optimal with respect to the choice of control effort penalty, performance penalty, dynamic compensator, and location of microphones and speakers) were then used to close the loop. For greater detail on the design of optimal, output feedback controllers, the reader is referred to the work of Toivonen (1985). Results from these control experiments are summarized below.

1. SISO feedback control

A schematic diagram of the fuselage and the control source, control error microphone, and an auxiliary microphone is illustrated in Fig. 8. For the purpose of this control system design, a single microphone was positioned 1 in. from the control source at location *d*. The disturbance was positioned at frame station 155 in. on the floor to minimize the effects of direct radiation on the measured control system performance. The loop transfer function for this example was presented previously in Fig. 5, and the response at the output was dominated by the direct field associated with the control source. Thus, for the purpose of observing the low-frequency modal characteristics of the controlled sound field, the frequency response between the disturbance and the control output was measured both before and after closing the loop. The measured frequency responses displayed in the following results are normalized to the reference pressure in air, $20E$-6 Pa, and A-weighting was applied to shape the fre-

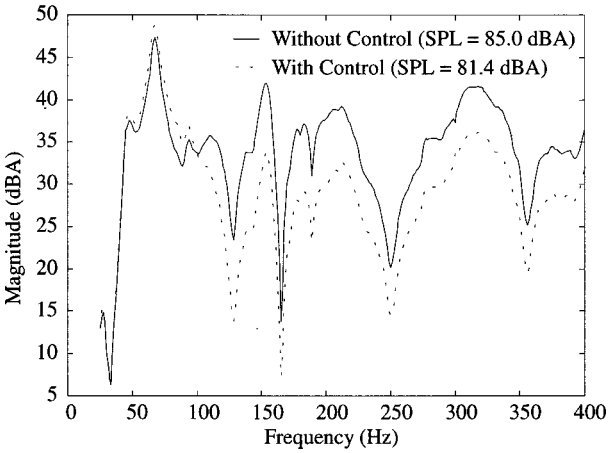


FIG. 9. The SISO frequency response from disturbance source to control output.

quency response. Thus, the magnitude of the frequency response is plotted in dBA normalized to the input voltage applied.

As illustrated in Fig. 9, a significant reduction, approximately 10 dBA, in the acoustic response at the control output between 100 and 400 Hz occurred upon closing the loop. The effective bandwidth of the control system extended from 100 Hz to above 400 Hz. However, as illustrated, the effect of the control system simply served to reduce the closed-loop magnitude. Cumulating the total power over the bandwidth, the control system results in approximately 3.6 dBA of attenuation. As detailed in Sec. III A, the reduction in magnitude was expected since the closed-loop gain is inversely proportional to the scalar feedback gain and the poles remain essentially unchanged.

To further demonstrate that the effect of the control system is local as opposed to global, the frequency response between the disturbance source and an auxiliary microphone positioned (frame station 265, circumferentially at 45 in., and radially inboard at 5 in.) in the fuselage was measured both before and after control and is presented in Fig. 10. As illustrated, the frequency response is unaffected by the control system, indicating that the SISO controller was limited to a local effect. In addition, the cumulative A-weighted power at

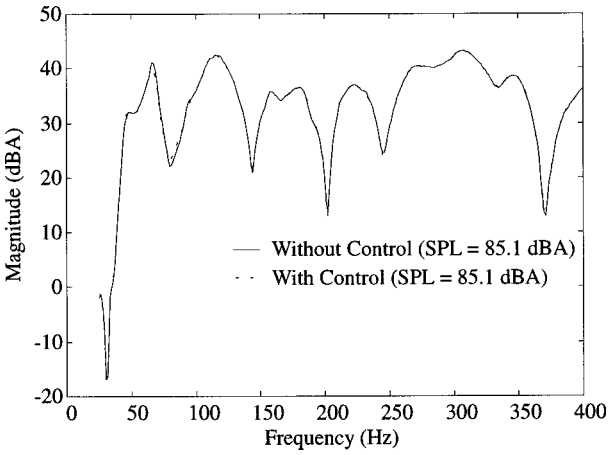


FIG. 10. Frequency response from disturbance to auxiliary microphone.

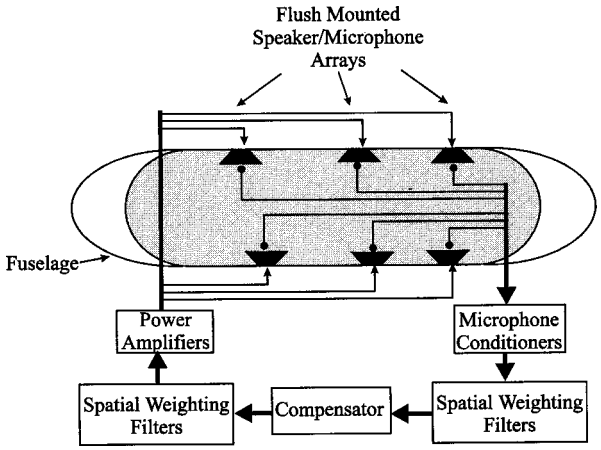


FIG. 11. Schematic diagram of fuselage configured with transducer array.

this microphone was the same before and after closing the loop.

A compensator was introduced into the loop transfer function for the SISO system to account for the transduction device dynamics, and while the effective bandwidth of the SISO system can be modified through loop shaping, the overall result remained the same: the control system serves to modify the sound field locally, but the input impedance of the reverberant plant remains relatively unaffected.

2. MIMO feedback control

Following the arguments made in Sec. III B, a substantially collocated array of transducers was implemented along the interior of the fuselage as illustrated in Fig. 11. The distribution of speakers and microphones was constrained physically by the mechanical construction of the airframe. However, the speakers were positioned so as to maximize control authority over a band of low-frequency acoustic modes. The outputs of the microphones and inputs to the speakers were ''weighted'' with a digital signal processor and summed to provide a distributed transducer capable of emphasizing the control of select acoustic modes and thereby separating the poles and zeros of the system transfer function as synthesized through the distributed transducer arrays. For this experiment, the transducer array was weighted in accordance with the second and third acoustic modes of the enclosure having resonant frequencies of 46 and 71 Hz, respectively.

The loop transfer function for an array weighted to emphasize the second and third acoustic modes of the enclosure at 71 Hz is illustrated in Fig. 12. As indicated in Fig. 12, the resonance (complex conjugate poles) and antiresonance (complex conjugate zeros) are well separated at 71 Hz.

Upon closing the loop, the frequency response between the disturbance and an auxiliary microphone not used in the array was measured. The closed loop frequency response from the control to output could not be obtained because the distributed transducer array was synthesized internally on a digital signal processor. However, for global control, the effects of the closed-loop response on the poles can be viewed through any system transfer function. The A-weighted frequency responses between the disturbance and auxiliary mi-

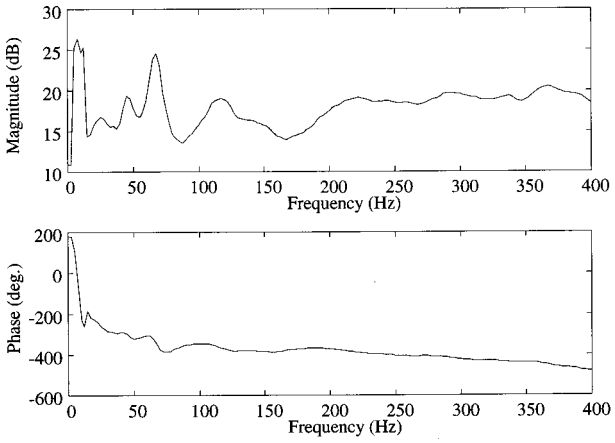


FIG. 12. Loop transfer function for weighted array.

crophone before and after closing the loop are presented in Fig. 13. As illustrated, a 6-dB reduction in the acoustic response of the third mode occurred. In addition, the energy dissipated in the second mode increased, resulting in approximately 3 dB of attenuation in the acoustic response of that mode. The acoustic response at alternative microphones reflected similar levels of dissipation, and through the use of a curve-fitting algorithm, an increase in the damping ratio from 5% to 10% (proportional to the attenuation) was obtained. Thus, the distributed, collocated array, control system resulted in dissipation of energy at the desired second and third acoustic modes of the fuselage.

However, as indicated by the cumulative, A-weighted, sound pressure level, the closed-loop system did little to alter the band power. The cumulative SPL before and after control remained relatively constant. The acoustic response of the enclosure in the bandwidth of the acoustic modes targeted in this investigation was between 10 and 20 dB below the response over the bandwidth of the disturbance. This offers some explanation for the lack of subjective audible difference in the sound field upon closing the loop. However, even upon limiting the bandwidth further to that of the modes under consideration, the system performance was essentially inaudible.

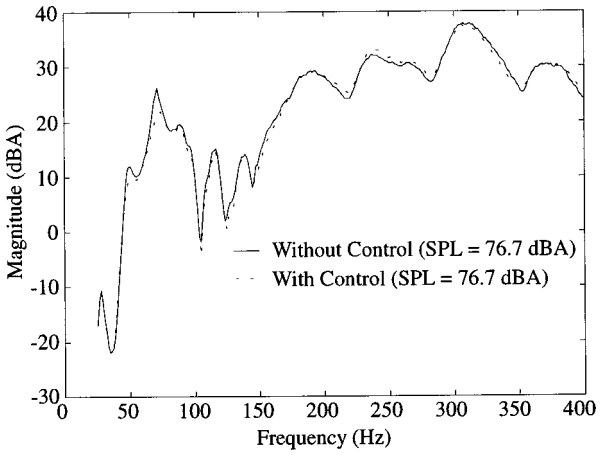


FIG. 13. Disturbance to auxiliary microphone frequency responses.

3. Summary

As expected, the SISO controller resulted only in local reductions in sound pressure levels, at the microphone used as the sensor. However, the MIMO controller resulted in more global levels of reduction at low frequency. In general, the MIMO performance was limited to very low bandwidth. To increase the bandwidth of reduction and the level of attenuation, a larger array of control sources (speakers) and sensors (microphones) would be required. While this might sound like a negative, one could actually decrease the size (diameter) of the speakers if the number used were increased. The primary objective is to selectively couple to the low-frequency modes of the enclosure, which will reduce spillover into higher-order modes. Upon accomplishing this task, much less control energy is required, and the collective array can be used to excite the low-frequency response of the enclosure as opposed to a single speaker with limited volume velocity. Additionally, the use of a volume velocity source would also enhance closed loop performance in such applications as detailed by Lane and Clark (1997).

IV. CONCLUSIONS

Results presented in Sec. III demonstrate that local feedback control loops with a single transducer pair will lead to a local quiet zone. The performance of the local control system is mandated by the modal density of the system, which is typically very high for a three-dimensional enclosure such as the fuselage of a typical business jet. Approximately 10 dB of attenuation was obtained locally over a bandwidth extending from 100 to 400 Hz. However, for global dissipation of reverberant energy, the system serves little purpose as it effectively operates on a infinite number of modes simultaneously, having very little effect on any single mode.

To alleviate this problem, a distributed array of transducers was implemented to emphasize the control of the second and third acoustic modes of the enclosure based upon knowledge of the spatial variation in the acoustic modes of interest. The MIMO array of transducers was weighted on the inputs and outputs with the same constant gain weighting functions such that a distributed SISO, substantially collocated, transducer pair was realized. This distributed transducer array served to emphasize the residues of the modes of interest, and thus more of the control effort was extended on these modes over the bandwidth of interest. As indicated in the results section, as much as 6 dB of attenuation was achieved at each of the acoustic modes of interest upon closing the loop with the distributed array control system implementation. Considering that the enclosure was reasonably well damped by the interior trim, this was a significant accomplishment, taking the damping ratio from approximately 5% to 10% at each acoustic mode.

The disappointing result was that the closed loop control of sound within the enclosure was subjectively inaudible. As indicated by the cumulative A-weighted power, 6 dB of attenuation at the third acoustic mode and 3 dB of attenuation at the second acoustic mode leads to effectively no difference in the summed power spectra at any particular microphone location. Thus, for the biological transducers, the ears,

the accomplishment was inaudible. Future efforts should be devoted to determining the bandwidth and characteristics of the disturbance for which active control of sound will have the greatest impact.

ACKNOWLEDGMENTS

The authors would like to acknowledge Dan Cole and Professor William Saunders for their contributions to this effort. In addition, the authors would like to thank Professor Chris Fuller for the use of laboratory facilities and equipment.

Balas, M. J. (**1979**). ''Direct Velocity Feedback Control of Large Space Structures,'' J. Guid. Control. Dyn. **2**(3), 252–253.

Bisplinghoff, R. L., and Ashley, H. (**1975**). *Principles of Aeroelasticity* (Dover, New York, 1975).

Clark, R. L. (**1997**). ''Accounting for out-of bandwidth modes in the assumed modes approach: Implications on colocated output feedback control,'' J. Dyn. Syst., Meas., Control. **119**(3), 390–395.

Clark, R. L., and Cole, D. G. (**1995**). ''Active Damping of Enclosed Sound Fields through Direct Rate Feedback Control,'' J. Acoust. Soc. Am. **97**, 1710–1716.

Clark, R. L., Cole, D. G., and Frampton, K. D. (**1996**). ''Phase Compensation for Feedback Control of Enclosed Sound Fields,'' J. Sound Vib. **195**(5), 701–718.

Kuttruff, H. (**1979**). *Room Acoustics* (Applied Science, Essex, England).

Lane, S. A., and Clark, R. L. (**1998**). ''Improving loudspeaker performance for active noise control applications,'' J. Audio Engin. Soc. **46**(6), 508–519.

MacMartin, D. G., and Hall, S. R. (**1991**). ''Structural control experiments using an Hinfinity power flow approach,'' J. Sound Vib. **148**(2), 223–241.

Meirovitch, L., and Baruh, H. (**1980**). ''Control of Self-Adjoint Distributed-Parameter systems,'' J. Guid. Control. Dyn. **5**(1), 60–66.

Nelson, P. A., and Elliott, S. J. (**1992**). *Active Control of Sound* (Academic, New York).

Olson, H. F., and May, E. G. (**1953**). ''Electronic sound absorber,'' J. Acoust. Soc. Am. **25**, 1130–1136.

Pierce, A. D. (**1981**). *Acoustics* (McGraw-Hill, New York).

Toivonen, H. T. (**1985**). ''A Globally Convergent Algorithm for the Optimal Constant Output Feedback Problem,'' Int. J. Control **41**, 1589–1599.

Reduction of the performance of a noise screen due to screen-induced wind-speed gradients. Numerical computations and wind-tunnel experiments

Erik M. Salomons
TNO Institute of Applied Physics, P.O. Box 155, 2600 AD Delft, The Netherlands

(Received 8 September 1998; revised 6 January 1999; accepted 16 January 1999)

Downwind sound propagation over a noise screen is investigated by numerical computations and scale model experiments in a wind tunnel. For the computations, the parabolic equation method is used, with a range-dependent sound-speed profile based on wind-speed profiles measured in the wind tunnel and wind-speed profiles computed with computational fluid dynamics (CFD). It is found that large screen-induced wind-speed gradients in the region behind the screen are responsible for a considerable reduction of the performance of the screen, for receivers near the boundary of the shadow region behind the screen. The screen-induced wind-speed gradients cause a considerable reduction of the size of the shadow region. If the screen-induced wind-speed gradients are taken into account, computed sound-pressure levels near the shadow boundary are in reasonable agreement with levels measured in the wind tunnel. In contrast, computed levels are considerably lower, up to 10 dB, if the screen-induced wind-speed gradients are ignored. This implies that the performance of a screen can be considerably improved if the screen-induced wind-speed gradients can be suppressed, e.g., by the use of ''vented'' screens. © *1999 Acoustical Society of America.* [S0001-4966(99)02504-7]

PACS numbers: 43.50.Gf, 43.50.Vt [MRS]

INTRODUCTION

Noise barriers are widely used for reducing noise from various sources in the open air, such as road and rail traffic, but also airplanes during ground tests. In general, measured barrier attenuations tend to be lower than predicted attenuations. The overestimation of barrier attenuation by computational schemes (e.g., the ISO 9613-2 scheme[1]) can be attributed to complex meteorological effects.

Meteorology affects sound propagation by the phenomenon of atmospheric refraction. Atmospheric refraction originates from wind-speed and temperature gradients. In many practical situations, the effect of the wind dominates, and temperature effects will be ignored in the following. This implies that sound waves propagating over the ground in downwind directions are refracted downwards, i.e., toward the ground. As a consequence, the performance of a noise screen between a source and a downwind receiver (see Fig. 1) is reduced by downward refraction.[2]

In situations without barriers, vertical wind-speed gradients are large near the ground and decrease with increasing height (here, we consider a wind-speed profile averaged over a few min; rapid turbulent fluctuations are averaged out). A screen induces additional wind-speed gradients, in particular in the region behind the barrier (see Fig. 1). These screen-induced wind-speed gradients enhance the downward refraction of sound waves. This effect will be referred to by the acronym RESWING=REfraction by Screen-induced WINd speed Gradients. The RESWING effect causes a reduction of screen attenuation (i.e., the attenuation of sound waves by a screen). In this paper, it will be shown that this reduction may be as large as 10 dB(A). ''Smoother'' noise barriers, such as an earth berm, induce smaller wind-speed gradients than a thin screen does, and therefore have a smaller RESWING effect.

It should be noted that the reduction of screen attenuation by RESWING is largest for receivers near the boundary of the shadow region behind a screen. In fact, RESWING causes a considerable reduction of the size of the shadow region. For receivers far from the shadow boundary, the effect of RESWING on the sound-pressure level is small.

Another complex atmospheric effect that reduces barrier attenuation is atmospheric turbulence. The mechanism that causes this reduction is scattering of sound waves by turbulent inhomogeneities in the atmosphere. Scattered waves enhance the sound-pressure level at a receiver in the shadow of a barrier.[3,4]

There are several methods to study the influence of wind and turbulence on the performance of a noise barrier:

(i) outdoor measurements on a full-scale barrier,[5,6]
(ii) measurements on a scale model in a wind tunnel,[7–10]
(iii) numerical computations.[4,9–11]

In this paper, we combine results of numerical computations and wind-tunnel experiments to study the influence of screen-induced wind-speed gradients on the performance of a screen.

For the numerical computations, we used the parabolic equation (PE) method.[12] Sound propagation over a screen in a refracting atmosphere can be computed with the PE method.[11] Complex meteorological effects such as turbulence[4,13,14] or screen-induced wind-speed gradients can be taken into account (approximately) with the PE method. Obviously, this requires an accurate description of the wind velocity field around the barrier. The description may be

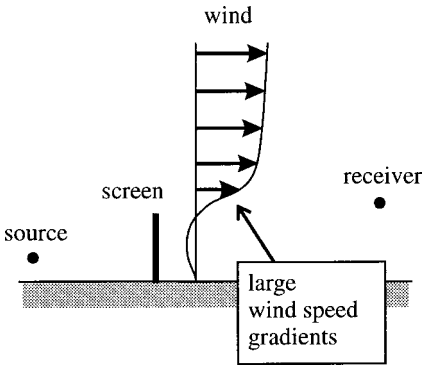


FIG. 1. Situation with a noise screen between a source and a downwind receiver. Screen-induced wind-speed gradients are indicated.

based on measured wind-speed profiles and/or computational fluid dynamics (CFD) results.[15]

For the wind-tunnel experiments, we used the setup shown schematically in Fig. 2. An important experimental problem is the elimination of reflections from the walls of the tunnel. In this work, we solved this problem by using a spark source.[7,8] A spark source produces a short sound pulse, and the reflections are easily eliminated from the received signal. Alternately, one may use a continuous source with a sweep signal, and compute the impulse response by signal processing.[16]

I. EXPERIMENTS IN A WIND TUNNEL

A. Experimental procedure

This section describes the measurements on an acoustic scale model in a wind tunnel. The measurements were performed in the atmospheric boundary layer wind tunnel of TNO-MEP in Apeldoorn, The Netherlands. The tunnel is 3 m wide, 2 m high, and about 25 m long. For the acoustic scale-model measurements, we used a section with a length of 3 m. Figure 3 shows the experimental setup, including the positions of the spark source, the microphones, and the noise screen. Table I gives the positions of the source, microphones, and the screen top, using the xyz coordinate system shown in Fig. 2. Microphone 1 near the source was used to check the reproducibility of the emitted signal. Microphones 2 and 3 at 3 m from the source were used to study the performance of the noise screen. The ground surface in the measurement section may be considered as rigid and smooth. The ground surface in the section upwind of the source was

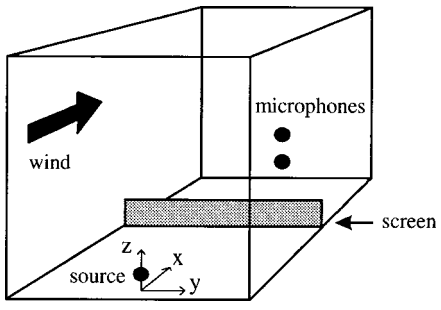


FIG. 2. Basic setup of the measurements on an acoustic scale model in a wind tunnel.

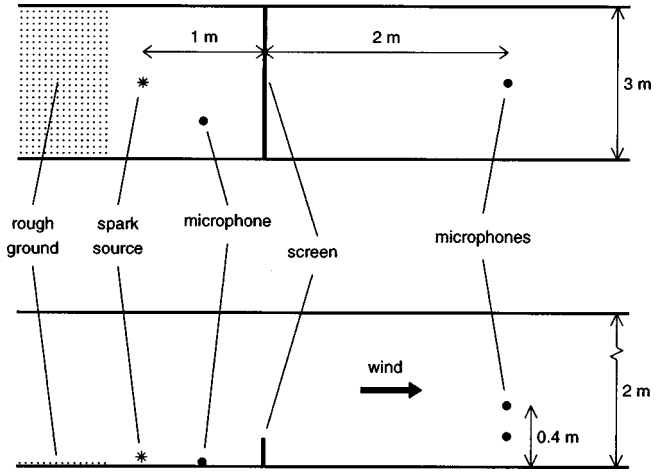


FIG. 3. Top view (top) and side view (bottom) of the setup in the wind tunnel.

covered with a rough layer, in order to create a realistic atmospheric wind-speed profile. The noise screen consisted of a thin wooden plate. As an acoustic barrier, the screen may be considered infinitely thin and rigid.

The spark source produced a short sound pulse, with a duration of about 0.1 ms. A small wind screen was used to eliminate effects of the wind on the spark signal. The peak pressure at microphone 1 was about 100 Pa, so we can assume that nonlinear acoustic effects are small. This assumption was verified by a few measurements with a spark source with 30 times lower peak pressures. Time signals were recorded with a sample frequency of 130 kHz. Analysis of the digital signals was performed on a computer. First, all reflections from the walls of the wind tunnel (not including the ground surface) were eliminated from the time signals. Next, spectra were computed of the sound exposure level $L_x = 10\lg(\langle p^2(t)\rangle_x / p_{\mathrm{ref}}^2)$, with $\langle p^2(t)\rangle_x$ the squared sound pressure averaged over a period of 1 s, and $p_{\mathrm{ref}} = 2\times10^{-6}$ Pa.

During the acoustic measurements, the wind speed in the tunnel was measured with a Pitot tube located about one meter upwind of the source, at a height of 0.30 m. This wind speed will be referred to as u_P. Acoustic measurements were performed for $u_P = 0,2,4,...,14$ m/s. Further, vertical profiles of the wind speed were measured in the tunnel for $u_P = 4$ and 10 m/s, using a hot wire anemometer. Wind speeds were averaged over a period of typically 1 min. Profiles were measured both with and without the screen. During the acoustic measurements, the temperature and relative humidity in the tunnel were recorded, which are the quantities that determine

TABLE I. Positions of the source, microphones, and the screen top, using the xyz coordinate system shown in Fig. 2.

	Scale model			Full-scale system		
	x (m)	y (m)	z (m)	x (m)	y (m)	z (m)
Spark source	0	0	0.063	0	0	1.89
Microphone 1	0.5	0.5	0	15	15	0
Microphone 2	3	0	0.20	90	0	6
Microphone 3	3	0	0.40	90	0	12
Screen top	1	⋯	0.185	30	⋯	5.55

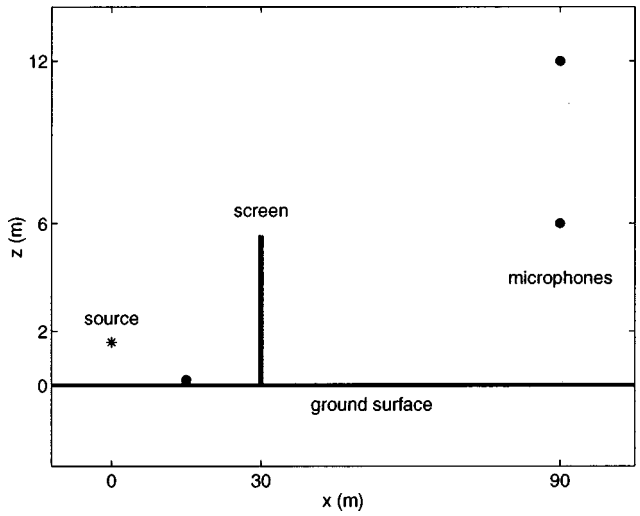


FIG. 4. Schematic representation of the full-scale system (scale 1:30), with x and z coordinate axes.

atmospheric attenuation.[17] Temperature gradients in the tunnel were negligible.

B. Scaling procedure

We assume a scaling ratio of 1:30, and use the following scaling rules:

(i) distances (including wavelengths) are multiplied by a factor of 30,
(ii) wind speeds are unaffected by scaling.

These rules give the best acoustic correspondence between the scale model and the full-scale system. Figure 4 shows the full-scale geometry. Table I gives the full-scale positions of the source, microphones, and the screen top. The first scaling rule implies that full-scale frequencies are a factor of 30 smaller than model frequencies. The second scaling rule implies that atmospheric refraction scales properly, as refraction depends on the gradient of the Mach number u/c (this follows from Snell's law). Since the sound speed c in the scale model is equal to the sound speed in the full-scale system, no scaling should be applied to the wind speed. This implies that the Reynolds number is not unaffected by the scaling, as usual in wind tunnel experiments;[15] this is of minor importance, however, since this work focuses on the effects of mean wind-speed profiles rather than the effects of turbulence.

In the following, all distances are full-scale distances and all frequencies are full-scale frequencies. The wind speed u_P measured upwind at a height of 0.30 m will be referred to as u_9, the wind speed at a height of 9 m, for the full-scale system.

C. Experimental results

After scaling of the frequencies, narrow-band spectra of the spark signals were converted to octave and third-octave band spectra. The highest frequency band that could be analyzed was the 1250-Hz band, the lowest was the 200-Hz band. Below 200 Hz, background noise levels were too high.

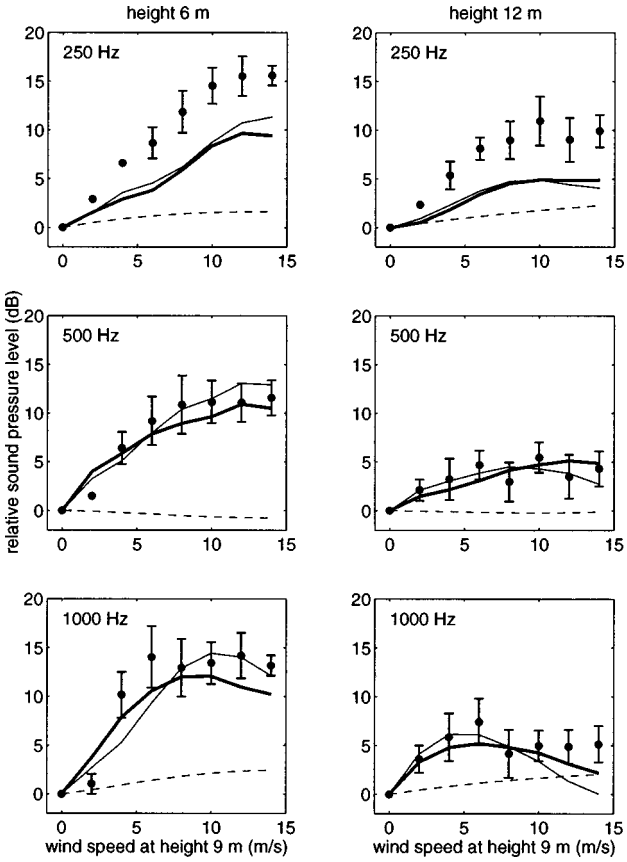


FIG. 5. Relative sound-pressure level of the spark source at microphone 2 at 6 m height (left column) and microphone 3 at 12 m height (right column) as a function of the wind speed u_9, for the octave bands 250, 500, and 1000 Hz (here, relative sound-pressure level=level minus level for zero wind speed). The measurements are represented by dots, with error bars indicating 95%-confidence intervals. Lines represent computations for profile 1 (dashed line), profile 2 (thin solid line), and profile 3 (thick solid line) (see Sec. II).

Figure 5 shows the measured relative sound-pressure level at the two microphones 2 and 3, as a function of wind speed u_9, for the octave bands 250, 500, and 1000 Hz. Each level represents a spectral average $10\lg\langle 10^{L_j(f)/10}\rangle$ of a series of ten spark signals ($j=1,2,...,10$). The error bars represent the 95%-confidence interval around this average. In particular at microphone 2 at 6 m height, the relative sound-pressure level increases strongly with increasing wind speed. In Sec. II, it will be shown that this increase is largely due to screen-induced wind-speed gradients.

It should be noted that the performance of a noise screen is best represented by the insertion loss, i.e., the difference in received sound-pressure level between the situations without and with the screen. This study, however, focuses on the *change* in received sound-pressure level due to screen-induced wind-speed gradients. This change in received sound-pressure level is equal to the change in insertion loss (except for a minus sign). To study changes in sound-pressure level, it is sufficient to consider only relative sound-pressure levels, as in Fig. 5. An important experimental advantage of considering the relative sound-pressure level as a function of wind speed is that corrections for the directivities of the source and the microphones are not necessary, as sound propagation is dominated by a single sound path, the path *source→screen top→receiver* (other sound paths with

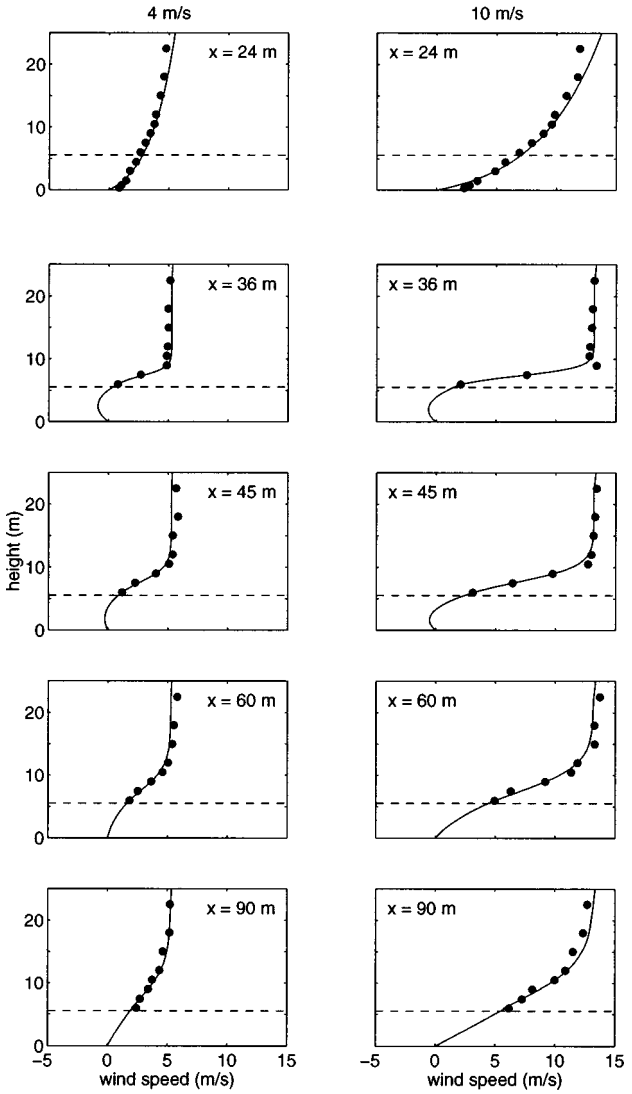


FIG. 6. Measured wind-speed profiles (dots), and wind-speed profile 3 used for the computations (solid curves), for $x=24$, 36, 45, 60, and 90 m. The left column is for $u_9=4$ m/s, the right column is for $u_9=10$ m/s. The dashed lines represent the screen height.

ground reflections are considerably weaker, in the present situation). In contrast, measuring the insertion loss would require directivity corrections. We note further that corrections for changes in atmospheric attenuation[17] turned out to be negligible.

Wind-speed profiles were measured both with and without the screen. In the situation without the screen, it was found that the profiles are well represented by the function[18] $u(z)=b\ln(1+z/z_0)$, with $b=u_9/\ln(1+z_9/z_0)$, $z_0=0.001$ m and $z_9=9$ m. The roughness length $z_0=0.001$ m represents an average of the two roughness lengths of the rough and smooth ground sections of the tunnel (see Fig. 3). A slightly more accurate fit of the data was obtained with a profile with two different logarithmic branches, which is sometimes used for the internal boundary layer originating from a change in ground roughness.[18] In the situation with the screen, profiles were measured for $x=24$, 36, 45, 60, and 90 m. The results are shown in Fig. 6. No data are shown in the figure for $x>30$ m and heights below the screen height of 5.55 m, as recirculation was observed for these heights and a hot wire anemometer cannot be used in a recirculating flow.[19] For heights $z\geqslant 6$ m, a stable flow in positive x direction was observed, so that the wind speeds shown in Fig. 6 represent horizontal wind vectors, in good approximation. The figure shows that wind-speed gradients are large for $x>30$ m and $6<z<10$ m. For example, the gradient for $z=8$ m and $u_9=4$ m/s is 2.0, 1.0, 0.6, and 0.4 s^{-1} at $x=36$, 45, 60, and 90 m, respectively, while the gradient in the situation without the screen is 0.05 s^{-1} in this case. With increasing height, the profile relaxes to a logarithmic profile $u(z)=b\ln(1+z/z_0)$, with $b=1.2\,u_9/\ln(1+z_9/z_0)$, $z_0=0.001$ m and $z_9=9$ m. The factor 1.2 accounts for the fact that the screen reduces the cross-sectional area of the tunnel, and consequently enhances the flow speed in the measurement section of the tunnel.

II. NUMERICAL COMPUTATIONS

This section describes numerical computations of the sound-pressure field in the system studied in the previous section (see Fig. 4). For these computations, we need the complete sound-speed field in the system, or equivalently, the complete wind-speed field (as a function of u_9). Section II A describes the construction of the wind-speed field, based on the measured wind-speed profiles and CFD computations. Section II B describes the acoustic computations and presents the computational results.

A. Construction of the wind-speed field

For the construction of the wind-speed field, we used the measured wind-speed profiles and results of CFD computations. The CFD computations were performed with the software package FLUENT,[20] for the same system as studied in the wind tunnel. FLUENT allows different turbulence models to be used, the $k\varepsilon$ model and the second-order Reynolds stress model.[18] It was found that the Reynolds stress model yields the best agreement with the measured wind-speed profiles. However, the agreement was not sufficient to allow the computed field to be used directly for the acoustic computations. The wind-speed field used for the acoustic computations was constructed from the measured wind-speed profiles, and this construction is described below. The computed field was used only as an aid for the extrapolation of the measured wind-speed profiles. For example, the computed field clearly shows that the effect of the screen on the flow field is largely concentrated in the region downwind of the screen; the effect on the field upwind of the screen is considerably smaller. Further, we verified that the scaling rules given in Sec. I B are satisfied, by repeating the computations for the full-scale system; only minor differences were found between the field in the model system and the field in the full-scale system. The computations also showed that vertical wind speeds are typically one order of magnitude smaller than horizontal wind speeds, except close to the screen, where the difference is smaller.

The wind-speed field is represented by the function $u=u(x,z,u_9)$. We distinguish three types of profiles (i.e., functions):

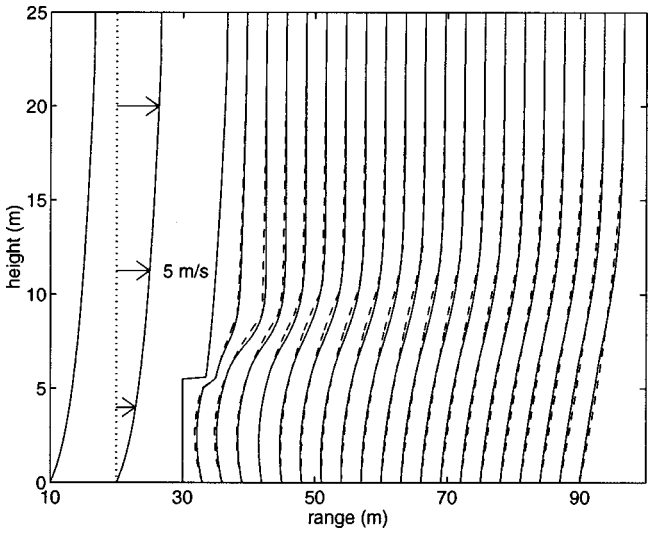


FIG. 7. Wind-speed profile 2 (dashed curves) and profile 3 (solid curves) at $x=10,20,30,33,36,...,90$ m, for $u_9=5$ m/s. The wind-speed scale is indicated for $x=20$ m, $z=11.25$ m.

(i) profile 1, the ''undisturbed'' profile $u(z)=b\ln(1+z/z_0)$, with $b=u_9/\ln(1+z_9/z_0)$, $z_0=0.001$ m and $z_9=9$ m.

(ii) profiles 2 and 3, based on two slightly different mathematical fits of the measured profiles.

Profiles 2 and 3 are represented in Fig. 7 (for $u_9=5$ m/s). In Fig. 6, profile 3 is compared with the measured profiles.

We now describe the mathematical representation of profiles 2 and 3. In the region $\{0\leqslant x\leqslant 90\text{ m}, z\geqslant 22.2\text{ m}\}$, both profiles are given by $u(z)=b\ln(1+z/z_0)$, with $b=1.2\,u_9/\ln(1+z_9/z_0)$, $z_0=0.001$ m and $z_9=9$ m (see Sec. I for the factor 1.2). In the region $\{0\leqslant x\leqslant 30\text{ m}, z\leqslant 22.2\text{ m}\}$, both profiles are given by $u(z)=b\ln(1+z/z_0)$, with $z_0=2$ m and b obtained from continuity at $z=22.2$ m. Here, we neglect any variation with x in the upwind field, as this is of minor importance in the present geometry (this was concluded from CFD computations and acoustic computations for different upwind fields). In the region $\{36\leqslant x\leqslant 90\text{ m}, z\leqslant 22.2\text{ m}\}$, profiles 2 and 3 are described below. In the region $\{30\leqslant x\leqslant 36\text{ m}, z\leqslant 22.2\text{ m}\}$, a smooth transition is made between the vertical profiles at $x=30\text{ m}+\varepsilon$ (with $\varepsilon\to 0$) and $x=36$ m, by linear interpolation in x. The vertical profile at $x=30\text{ m}+\varepsilon$ (i.e., just behind the screen) is equal to the profile at $x=30\text{ m}-\varepsilon$ for $z>5.55$ m, equal to zero for $z<5.00$ m, while linear interpolation in z is used for $5.00<z<5.55$ m.

We now describe the mathematical representation of profile 2 in the region $\{36\leqslant x\leqslant 90\text{ m}, z\leqslant 22.2\text{ m}\}$. For $x=36$, 45, 60, or 90 m and $u_9=4$ or 10 m/s, profile 2 for $z\leqslant 22.2$ m is given by $u=A_2z^2+A_1z$, with coefficients A_2 and A_1 given in Table II. For intermediate values of x, linear interpolation is used. For example, $u(70\text{ m},z,u_9)=\frac{2}{3}u(60\text{ m},z,u_9)+\frac{1}{3}u(90\text{ m},z,u_9)$. For intermediate values of $u_9<10$ m/s, linear interpolation is used (note that $u=0$ for $u_9=0$ m/s). For $u_9>10$ m/s, linear extrapolation is used.

Profile 3 in the region $\{36\leqslant x\leqslant 90\text{ m}, z\leqslant 22.2\text{ m}\}$ was constructed in a similar way as profile 2. First, the measured wind-speed profiles for $x=36$, 45, 60, and 90 m and $u_9=4$ and 10 m/s were fitted by smooth mathematical functions. As these functions are more complex than for profile 2, we do not give the functions for profile 3 but refer to Fig. 6, where the functions are shown together with the measured wind speeds. In the same way as for profile 2, linear interpolation is used for intermediate values of x and u_9, and linear extrapolation for $u_9>10$ m/s.

TABLE II. Coefficients A_2 and A_1 for profile 2.

	$u_9=4$ m/s		$u_9=10$ m/s	
x (m)	A_2 ($s^{-1}m^{-1}$)	A_1 (s^{-1})	A_2 ($s^{-1}m^{-1}$)	A_1 (s^{-1})
36	0.1378	−0.6993	0.3830	−1.9600
45	0.0650	−0.1968	0.1556	−0.4233
60	0.0195	0.1865	0.0280	0.6534
90	−0.0092	0.4543	−0.0268	1.1886

B. Computations of the sound-pressure field

Computations of the sound-pressure field in the system shown in Fig. 4 were performed with the Crank–Nicholson PE method.[12] The screen was taken into account by the method described in Ref. 11. Octave-band levels were computed by averaging over 30 frequencies per band. A grid spacing $d\leqslant 0.1\,\lambda$ was used, where λ is the wavelength. Spurious numerical oscillations[11] were eliminated by spatial averaging over a horizontal interval of 5 λ. A system height of 100 m was used, and an absorbing layer of about 50 λ on top of the system.

For the sound-speed field, we used the effective sound speed $c=c_0+u$, with $c_0=343$ m/s and $u=u(x,z,u_9)$ one of the three wind-speed profiles described in Sec. II A. The dependence of the effective sound speed on the range x was taken into account by changing the vertical profile $c(z)$ after each finite-difference range step.[12] It should be noted that recently a PE method for a moving atmosphere was developed.[21] For the present study, however, the PE method based on the effective sound-speed approximation was considered sufficiently accurate.

Computed values of the relative sound-pressure level of the spark source are shown in Fig. 5 as a function of the wind speed u_9. The variation of the sound-power level of the spark source within an octave band was taken into account by using the third-octave sound-power levels L_{Wx} for the octave band averaging of computational results ($L_{Wx}=96$, 99, 102, 106, 109, 114, 118, 118, and 113 dB for the third-octave bands 200 to 1250 Hz). The figure shows that the variation of the received sound-pressure level with wind speed is small for profile 1, without screen-induced wind-speed gradients. In contrast, the computed relative sound-pressure level for profiles 2 and 3, with screen-induced wind-speed gradients, increases with wind speed to about +10 dB for microphone 2, and about +5 dB for microphone 3. The computational results for profiles 2 and 3 are in reasonable agreement with the experimental data. The largest deviations occur for the 250-Hz band. Possible origins of the deviations are: (i) differences between the wind-speed field assumed for the computations and the real wind-speed profile in the wind

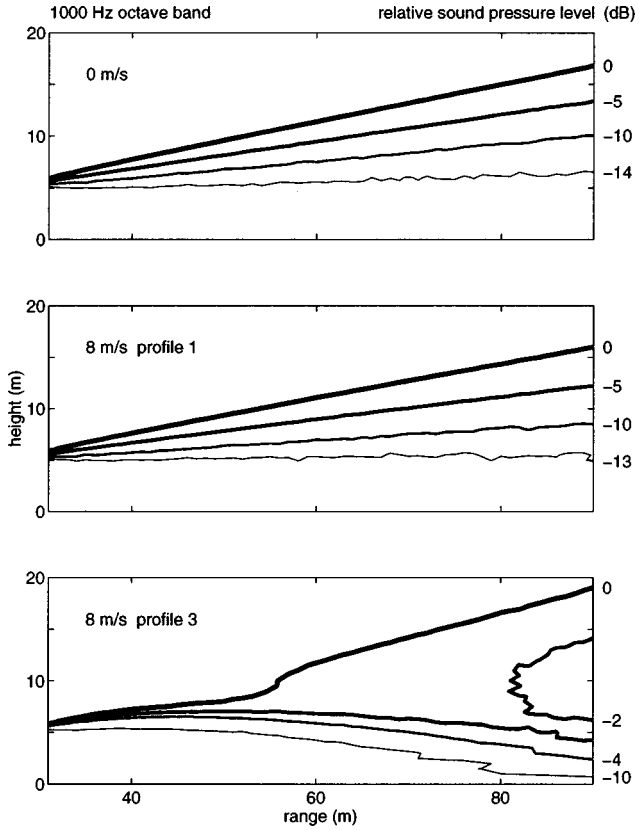


FIG. 8. Computed contours of the relative sound-pressure level (level minus free-field level) for the 1000-Hz octave band, for three wind-speed profiles: $u_9=0$ m/s (top), profile 1 for $u_9=8$ m/s (middle), profile 3 for $u_9=8$ m/s (bottom).

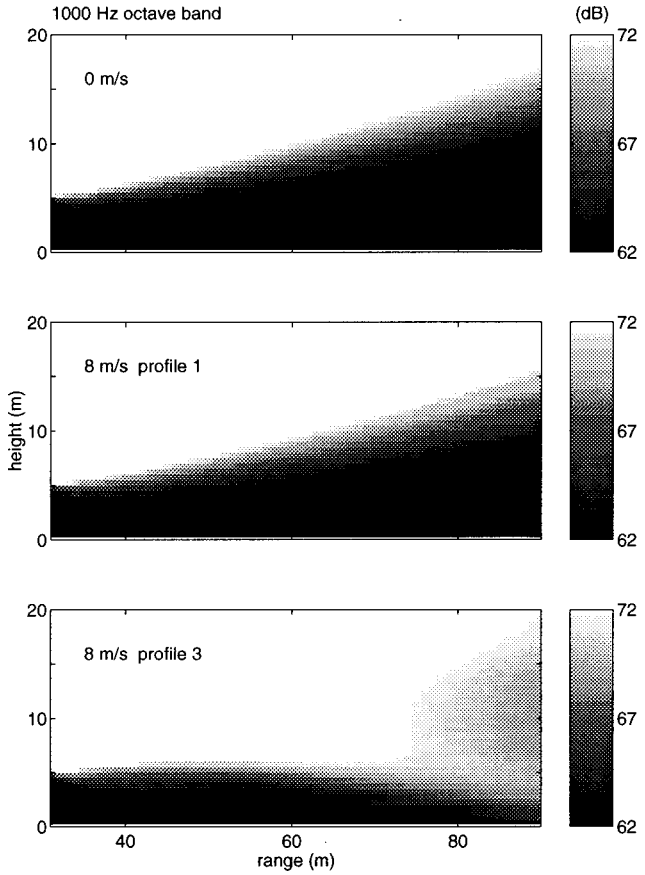


FIG. 9. Grey-level plots of the absolute sound-pressure level, corresponding with the contour plots shown in Fig. 8. Levels above 72 dB are set equal to 72 dB, levels below 62 dB are set equal to 62 dB.

tunnel (e.g., vertical wind components are ignored by the PE method), and (ii) the neglect of atmospheric turbulence in the computations. Figure 5 clearly shows, however, that the RESWING effect is a large one, and that this effect can be computed with the PE method with reasonable accuracy.

Figures 8 and 9 show graphical representations of the complete sound-pressure field behind the screen, for the 1000-Hz band. For the 500- and 250-Hz bands, we found similar results. Figure 8 shows contours of equal relative sound-pressure level, for three wind-speed profiles: (i) $u=0$ m/s, (ii) profile 1 for $u_9=8$ m/s, and (iii) profile 3 for $u_9=8$ m/s. Figure 9 shows corresponding grey-level plots of the absolute sound-pressure level, i.e., the sound-pressure level including geometrical attenuation and atmospheric attenuation[17] (for temperature 10 °C and relative humidity 80%). From these figures, one sees that the sound-pressure level is enhanced by screen-induced wind-speed gradients (profile 3) in a large region behind the screen. Figure 9 clearly shows that RESWING causes a considerable reduction of the size of the shadow region. As a consequence, the effect of RESWING on the sound-pressure level is large for receivers near the shadow boundary and small for receivers far from the shadow boundary (the expressions ''near the shadow boundary'' and ''far from the shadow boundary'' should be interpreted in terms of diffraction angles[22]). For example, from the sound-pressure fields shown in Fig. 9, we find that RESWING causes an increase of the sound-pressure level of 11 dB at receiver position $(x,z)=(90\,\text{m},6\,\text{m})$, 10 dB at (90 m, 2 m), 3 dB at (60 m, 2 m), and 1 dB at (40 m, 2 m).

III. IMPROVING THE DOWNWIND PERFORMANCE OF A SCREEN

As indicated before, the performance of a screen is best represented by the insertion loss. For zero wind speed, computed insertion loss spectra of the screen shown in Fig. 4 are given in Table III. With increasing wind speed, the insertion loss varies depending on the variations with wind speed of the received sound-pressure level in the situations with and without the screen. Figure 5 shows the variation in the situation with the screen. This figure also shows that the received sound-pressure level in the situation with the screen could be considerably reduced *if* the RESWING effect could be suppressed. In other words, the downwind insertion loss could be considerably enhanced if the RESWING effect could be suppressed.

To suppress the RESWING effect, one has to reduce the large wind-speed gradients in the region indicated in Fig. 1.

TABLE III. Computed octave-band levels of the insertion loss of the screen in Fig. 4, for zero wind speed.

	250 Hz (dB)	500 Hz (dB)	1000 Hz (dB)
Microphone 2	16	12	17
Microphone 3	6	9	10

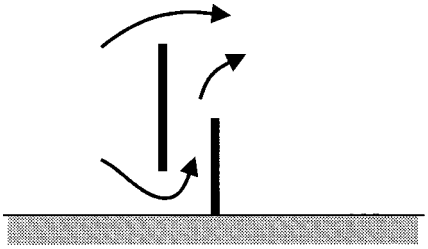


FIG. 10. Schematic representation of a vented screen, with air flowing over the screen and through the opening in the screen.

It should be emphasized that the sound path from the screen top to the receiver passes just through this region. This suggests that the RESWING effect may be suppressed by the use of a vented screen, i.e., a screen with openings for air flow. The openings should be constructed in such a way that sound waves propagating through the openings are strongly attenuated. An example of a vented screen is shown in Fig. 10. In this way, the large wind-speed gradients in the region indicated in Fig. 1 may be reduced at the expense of an increase of wind-speed gradients at heights well below the screen height. The latter gradients, however, do not affect the sound propagation over the screen. Another approach to suppress the RESWING effect may be the use of a screen with a top that is ''aerodynamically streamlined.'' An extreme example is an earth berm, but one may also think of a cylinder at the top of the screen.

IV. CONCLUSIONS

It has been found that the RESWING effect causes a large reduction of the downwind performance of a noise screen, for receivers near the shadow boundary. The size of the shadow region behind the screen is considerably reduced by RESWING. This result has been obtained by a combination of numerical computations and scale-model experiments in a wind tunnel. For the computations, a wind-speed field has been constructed that is based on wind-speed measurements in the wind tunnel and on CFD computations. The field contains large wind-speed gradients in the region behind the screen. Computed sound-pressure levels in the shadow of the screen are in reasonable agreement with measured levels. In contrast, computed levels near the shadow boundary are considerably lower if the large wind-speed gradients are ignored. This implies that the downwind performance of a noise screen can be considerably improved if the RESWING effect can be suppressed. The development of screens with a suppressed RESWING effect, such as vented screens, will be the subject of future work.

ACKNOWLEDGMENTS

The author is grateful to R.R. Parchen, J.C. Bruggeman, A.A.F.M. Beeks, J. van 't Hof, L.J.M. Hopmans, and J.D. van der Toorn for stimulating discussions and helpful comments on this work, and to G.P. Boerstoel for performing the CFD computations.

[1] ISO 9613-2, *Acoustics—Attenuation of Sound During Propagation Outdoors—Part 2: General Method of Calculation*, 1st ed., 1996-12-15.
[2] E. M. Salomons, ''Noise barriers in a refracting atmosphere,'' Appl. Acoust. **47**, 217–238 (1996).
[3] G. A. Daigle, ''Diffraction of sound by a noise barrier in the presence of atmospheric turbulence,'' J. Acoust. Soc. Am. **71**, 847–854 (1982).
[4] J. Forssen, ''Calculation of sound reduction by a screen in a turbulent atmosphere using the parabolic equation method,'' Acust. Acta Acust. **84**, 599–606 (1998).
[5] W. E. Scholes, A. C. Salvidge, and J. W. Sargent, ''Field performance of a noise barrier,'' J. Sound Vib. **16**, 627–642 (1971).
[6] D. J. Saunders and R. D. Ford, ''A study of the reduction of explosive impulses by finite sized barriers,'' J. Acoust. Soc. Am. **94**, 2859–2875 (1993).
[7] R. DeJong and E. Stusnick, ''Scale model studies of the effects of wind on acoustic barrier performance,'' Noise Control Eng. J. **6**, 101–109 (1976).
[8] K. B. Rasmussen, ''Model experiments related to outdoor propagation over an earth berm,'' J. Acoust. Soc. Am. **96**, 3617–3620 (1994).
[9] K. B. Rasmussen, ''Sound propagation over screened ground under upwind conditions,'' J. Acoust. Soc. Am. **100**, 3581–3586 (1996).
[10] N. Barriere and Y. Gabillet, ''Application of the GFPE method to sound propagation over a barrier with realistic wind gradients,'' Acust. Acta Acust. (submitted).
[11] E. M. Salomons, ''Diffraction by a screen in downwind sound propagation: a parabolic-equation approach,'' J. Acoust. Soc. Am. **95**, 3109–3117 (1994).
[12] K. E. Gilbert and M. J. White, ''Application of the parabolic equation to sound propagation in a refracting atmosphere,'' J. Acoust. Soc. Am. **85**, 630–637 (1989).
[13] 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).
[14] P. Chevret, Ph. 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).
[15] E. Simiu and R. H. Scanlan, *Wind Effects on Structures* (Wiley, New York, 1978), pp. 111–170, 318–342.
[16] Y. Gabillet, G. A. Daigle, and A. L'Esperance, ''Sound propagation in a wind tunnel: comparison of experiments with FFP and residue solution,'' Appl. Acoust. **43**, 321–331 (1994).
[17] ISO 9613-1, *Acoustics—Attenuation of Sound During Propagation Outdoors—Part I: Calculation of the Absorption of Sound by the Atmosphere*, 1st ed., 1993-06-01.
[18] R. B. Stull, *An Introduction to Boundary Layer Meteorology* (Kluwer, Dordrecht, 1988), pp. 376–378, 596–599.
[19] H. H. Bruun, *Hot-Wire Anemometry* (Oxford University Press, Oxford, 1995), pp. 234, 235.
[20] FLUENT, Computational Fluid Dynamics Software, version 4.3, January 1995, Fluent, Incorporated, Centerra Resource Park, 10 Cavendish Court, Lebanon, NH 03766.
[21] V. E. Ostashev, D. Juvé, and P. Blanc-Benon, ''Derivation of a wide angle parabolic equation for sound waves in inhomogeneous moving media,'' Acust. Acta Acust. **83**, 455–460 (1997).
[22] A. D. Pierce, *Acoustics. An Introduction to its Physical Principles and Applications* (American Institute of Physics, Woodbury, NY, 1991), p. 495.

Actively created quiet zones for broadband noise using multiple control sources and error microphones

Jingnan Guo[a] and Jie Pan
Department of Mechanical and Materials Engineering, University of Western Australia, Nedlands, WA 6907 Australia

(Received 7 August 1997; accepted for publication 29 November 1998)

Previous research on the quiet zones created in the free space, using multiple control sources and error microphones equally located in two parallel lines or in two parallel planes, demonstrated that there exists a range of optimal spacing for the control sources and error microphones. With the optimal arrangement, the largest area of quiet zone can be created with the least increase of total power output. This range for the optimal spacing is related to the configuration of the control system and the wavelength of the noise. The effectiveness of the optimally arranged control system to broadband noise is investigated in this paper. The results from both numerical simulation and experiment show that a control system optimally arranged for a particular frequency will be effective not only to the noise at the frequency, but also to that within a frequency band specified in this paper. © *1999 Acoustical Society of America.* [S0001-4966(99)04303-9]

PACS numbers: 43.50.Ki [MRS]

INTRODUCTION

The present authors have investigated the active control of a local sound field in free space, which attenuates the noise and creates the quiet zones in some desired places and is always referred to as the local control strategy. It was found that the performance of the control system is largely dependent on the configuration of the control system.[1–3] This research on the typical control systems with multiple control sources and error microphones equally placed in two parallel lines and in two parallel planes demonstrated that the spacings of the control sources and error microphones are vital to the performance of the control system. A range of optimal spacing of control sources and error microphones was discovered. When the control actuators and sensors are arranged within this range of spacing, the control system can create the largest area of quiet zone with the least increase in total power output.

As the optimal configuration of the control system is associated with the wavelength of the noise to be controlled, it is important to know if an optimally arranged control system for a specific frequency is also effective for other frequencies. Under practical circumstances, the control system usually has to deal with a broadband noise or a noise with multiple frequency components. The main objective of this paper is to investigate the application of the well-developed multiple active noise-control system to the broadband noise. The computer simulation results show that an optimally arranged multichannel control for a specific frequency is also optimal for a frequency band, which has been verified by the experiments in an anechoic chamber. The upper and lower limits of the effective frequency band have been analytically obtained. The results of this paper extend the optimal local control system from a single to a broadband frequency, and make the optimal control system more practical. The paper also provides the guideline in designing the optimal local control system for the broadband frequency noise. In this paper, the broadband noise refers to the noise with its frequency components covering a range of frequency.[4]

The two main indicators of the local control efficiency, total power-output increase, and the size of the quiet zone, are used in this investigation. The simulation analyses are to understand the relationship between the configuration of the control system and its performance for different frequency range, and to find out the effective frequency band. Both the primary noise source and the control sources are modeled as constant-volume velocity sources in the analyses. The minimization of the sound pressure with respect to the control-actuator output implies feed-forward control. The transient interaction between controller and physical system is included in this paper.

I. OPTIMAL CONFIGURATION OF THE SYSTEM

Two active noise-control systems are considered in this analysis. One consists of N control sources and N error microphones equally spaced in two parallel lines shown in Fig. 1, and the other has $N\times N$ control sources and $N\times N$ error microphones equally spaced in two parallel planes shown in

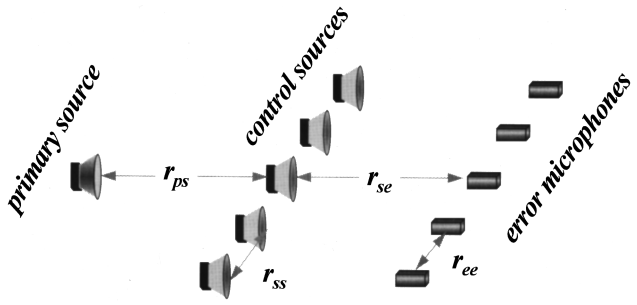


FIG. 1. The multichannel control system with control sources and error microphones equally spaced in two parallel lines.

[a]Present address: Occupational Hygiene Programme & Department of Mechanical Engineering, The University of British Columbia, 3rd Floor East Mall, Vancouver, British Columbia V6T 1Z3, Canada.

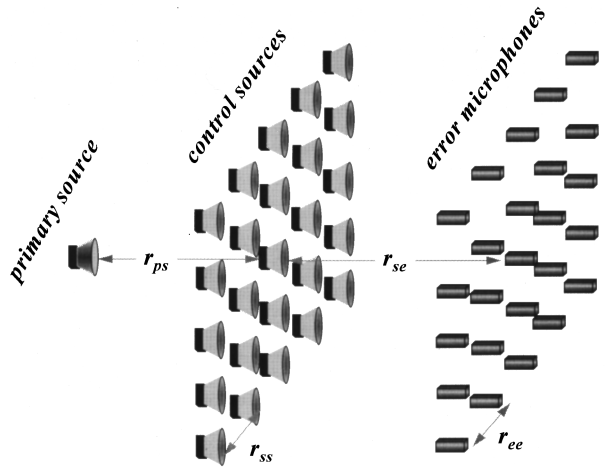


FIG. 2. The multichannel control system with control sources and error microphones equally spaced in two parallel planes.

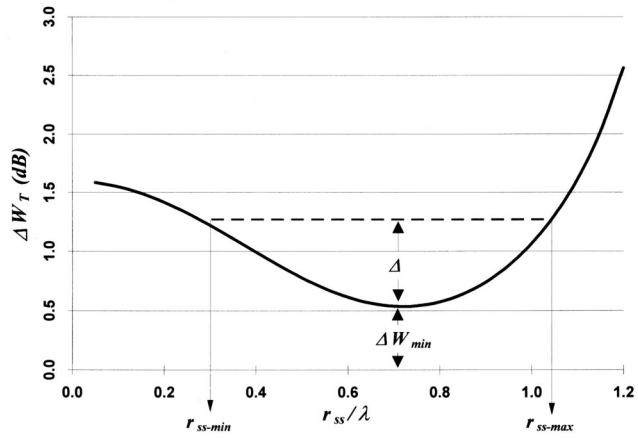


FIG. 3. The total power-output increase of the control system with respect to the spacing of the control sources and error microphones.

Fig. 2. The separation of the control sources and that of the microphones will be the same in the following discussions, i.e., $r_{ss}=r_{ee}$. The local control strategy is used, and the quiet zone can be created in some desired areas by the control system. Quiet zone is defined as the area where the primary sound-pressure level is attenuated by more than 10 dB.[5,6] The strengths of the N (or $N\times N$) control sources are adjusted so as to drive the total sound pressures at N (or $N\times N$) error microphones to a minimum, as[2,3]

$$\mathbf{q}_s=-\mathbf{Z}_{se}^{-1}\mathbf{Z}_{pe}q_p, \tag{1}$$

and the resultant sound pressure at any position of the space becomes

$$P_T=(Z_p-\mathbf{Z}_s\mathbf{Z}_{se}^{-1}\mathbf{Z}_{pe})q_p, \tag{2}$$

where q_p is the primary source strength, $\mathbf{Z}_{se}$ is the transfer-impedance matrix from the control sources to error mirophones, $\mathbf{Z}_{pe}$ is the column vector of transfer impedance between the primary source and the error microphones, $\mathbf{Z}_s$ is the row vector of transfer impedance from the control sources to a position in space, and Z_p is the transfer impedance from the primary source to the position where P_T is evaluated. The total radiated acoustic power of the system is given by[7,8]

$$W_T=\tfrac{1}{2}[Z_0|q_p|^2+\mathbf{q}_s^H\,\mathrm{Re}(\mathbf{Z}_{ss})\mathbf{q}_s+q_p^*\,\mathrm{Re}(\mathbf{Z}_{ps}^H)\mathbf{q}_s+\mathbf{q}_s^H\,\mathrm{Re}(\mathbf{Z}_{ps})q_p], \tag{3}$$

where $\mathbf{Z}_{ss}$ is the transfer-impedance matrix among the control sources, $\mathbf{Z}_{ps}$ is the column vector of transfer impedance from the primary source to control sources, $Z_0=\omega^2\rho_0/4\pi c_0$, and the superscript H denotes the Hermitian transpose. Substituting Eq. (1) into Eq. (3), the total power output of the control system can be rewritten as

$$W_T=W_0\{1+[(\mathbf{Z}_{se}^{-1}\mathbf{Z}_{pe})^H\mathrm{Re}(\mathbf{Z}_{ss})(\mathbf{Z}_{se}^{-1}\mathbf{Z}_{pe})-\mathrm{Re}(\mathrm{Re}(\mathbf{Z}_{ps}^H)(\mathbf{Z}_{se}^{-1}\mathbf{Z}_{pe}))]/Z_0\}, \tag{4}$$

where W_0 is the power output of the primary source when the control system is off. It has been demonstrated that for the local strategy, in which the main objective is to create the quiet zone in some desired area, the total sound-power output of the control system always increases after the controller is switched on, i.e., $W_T>W_0$. The optimal configurations of the control system are those that can create the largest area of quiet zone with least increase of total power output. It had been found that there always exist some configurations of the control system that have the low increase of total power output. Those configurations corresponding to the low increase of total power output can be seen as a range of spacing ($r_{ss\text{-}\min}\leqslant r_{ss}\leqslant r_{ss\text{-}\max}$) between adjacent control sources and error microphones in Fig. 3 as an example, where $\Delta W_T=10\log_{10}(W_T/W_0)$ is the total power increase after the control in dB. $\Delta=r_{se}/(r_{se}+r_{ps})$ in dB is the relative level derivation used to define the low power-output increase range.[2]

It has been found by the authors that the range with low power-output increase is the arrangement where adjacent control sources enhance each other at the error sensors; the control system then requires relatively small sound power to cancel the primary sound fields. It is a function of the distances between the primary and control sources r_{ps}, control-source array and error-microphone array r_{se}, and the wavelength of the noise λ; the upper limit of this range is determined as[2,3]

$$r_{ss\text{-}\max}\cong\begin{cases}\dfrac{\lambda}{2}\sqrt{1+\dfrac{4r_{se}}{N\lambda}}, & N=2,4,6,\ldots,\\[2ex] \dfrac{\lambda}{2}\sqrt{1+\dfrac{N+1}{N-1}\dfrac{4r_{se}}{N\lambda}}, & N=3,5,7,\ldots,\end{cases} \tag{5}$$

and the lower limit is given as

$$r_{ss\text{-min}} \cong \begin{cases} \dfrac{5\lambda}{2}\exp\left\{-\left[\dfrac{3(\lambda+0.04r_{ps})}{2r_{se}-\lambda}+\dfrac{20\lambda}{15\lambda+r_{ps}}\right]\right\}, & N=4,6,8,..., \\ \dfrac{3\lambda(N+1)}{N}\exp\left\{-\left[\dfrac{\lambda+2r_{ps}}{2(2r_{se}-\lambda)}+\dfrac{12\lambda}{5\lambda+r_{ps}}\right]\right\}, & N=3,5,7,..., \end{cases} \tag{6}$$

for the system with control sources and error microphones equally spaced in two parallel lines, or

$$r_{ss\text{-min}} \cong \begin{cases} \dfrac{1.4\lambda}{20}\left(\ln\left(\dfrac{r_{ps}}{\lambda}+\dfrac{1}{2}\right)+0.5\right)\ln\left(\dfrac{r_{se}}{\lambda}+\dfrac{1}{2}\right)+\dfrac{\lambda}{3}, & N=2,4,6,..., \\ \dfrac{(N+1)\lambda}{20(N-2)}\left(1.25\ln\left(\dfrac{r_{ps}}{\lambda}+\dfrac{1}{2}\right)+0.7\right)\ln\left(\dfrac{r_{se}}{\lambda}+\dfrac{N+1}{4(N-1)}\right)+\dfrac{\lambda}{3}, & N=3,5,7,..., \end{cases} \tag{7}$$

for the system with control sources and error microphones equally spaced in two parallel planes. The lower limit $r_{ss\text{-min}}$ for $N=2$ in Eq. (6) corresponds to a special case and has been presented in Ref. 2. It has also been found that the system can create the largest area of quiet zone when it is arranged in this range (i.e., $r_{ss\text{-min}} \leqslant r_{ss} \leqslant r_{ss\text{-max}}$); the sound pressure in the areas outside the quiet zone also has the least increase. Therefore, this range of low total power-output increase is also called the optimal range of the sensor and actuator spacing.

II. EFFICIENCY TO BROADBAND NOISE

In the previous work, the optimal configuration was discussed in terms of the wavelength of the noise to be controlled, and the control efficiency of the control system for the noise with that particular wavelength has been demonstrated. It should also be noted that the optimal configuration of the control system is given as a range. This implies that the optimally arranged control system for a particular frequency may also be effective to the noise within a frequency range.

A. Overlap of optimal ranges

As the optimal spacing of the control sources and error microphones has a range of selections, the optimal range for different frequency may overlap. In other words, there might exist a common range of optimal spacing for a frequency band. When it is arranged in the common range, the configuration of the control system is not only optimal to the noise at the specified frequency, but also to that in a frequency band. To illustrate this common range, Fig. 4 shows the optimal ranges of a control system corresponding to frequencies at 300, 400, and 500 Hz. The system consists of nine control sources and nine error microphones ($N=9$) equally spaced in two parallel lines; the distance from the primary source to the control-source array is 5 m, and the distance between the control-source array and error-microphone array is 10 m.

It can be seen that a common range of optimal spacing exists. This range includes r_{ss} from 0.87 to 1.02 m, which corresponds to the lower limit for noise at $f=300$ Hz and upper limit for that at $f=500$ Hz, respectively. This indicates that when the control system is arranged in this common optimal range, the control system is effective for a broadband noise with at least the frequency range (300–500 Hz).

B. Width of the effective-frequency band

The width of the effective-frequency band is studied for various control systems. A case study of a typical control system is used as an illustrating example. The control system has nine control sources and nine error microphones equally spaced in two parallel lines, and is optimally arranged to attenuate the noise of 500 Hz; the corresponding wavelength is 0.688 m. The distance from the primary source to the control-source array is $r_{ps}=5\lambda=3.44$ m, and the distance between the control-source array and error-microphone array

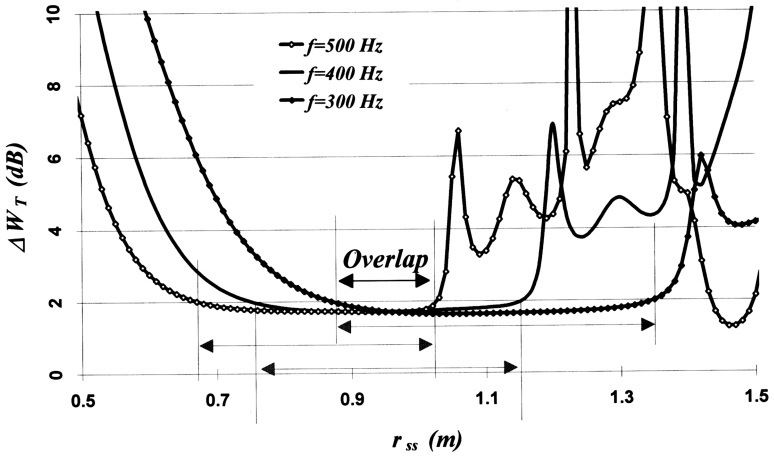


FIG. 4. Optimal range overlaps of a control system for different frequencies.

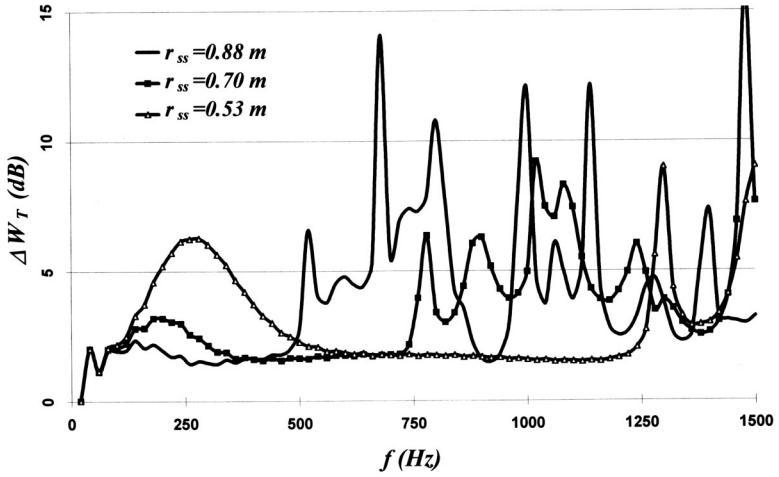


FIG. 5. The total power-output increase of the multichannel control system ($N=9$) with respect to the frequency of the noise for three spacings of control sources.

$r_{se}=10\lambda=6.88$ m. Using Eqs. (5) and (6), the optimal separation of the adjacent control sources can be selected from the range of $(0.75\lambda-1.28\lambda)$, which can be rewritten in meters as (0.53–0.88 m). The control efficiency of the system to the broadband noise is also analyzed in two aspects: the total power-output increase of the control system and the size of the quiet zone.

The total power-output increase of the control system is calculated in a wide frequency range (0–1500 Hz). Three spacings of the control sources within the optimal range for 500 Hz, $r_{ss}=0.53$, 0.70, and 0.88 m, are examined. They correspond to the lower and upper limits and the middle of the optimal range. The results are shown in Fig. 5.

When the lower limit of the optimal range $r_{ss}=0.53$ m is chosen as the spacing of the control sources, the least power-output increase lies in the range higher than 500 Hz (500–1220 Hz). If the upper limit $r_{ss}=0.88$ m is chosen as the spacing of the control sources, the control system has the least total power-output increase for the noises with the frequency lower than 500 Hz (220–500 Hz). For a spacing in

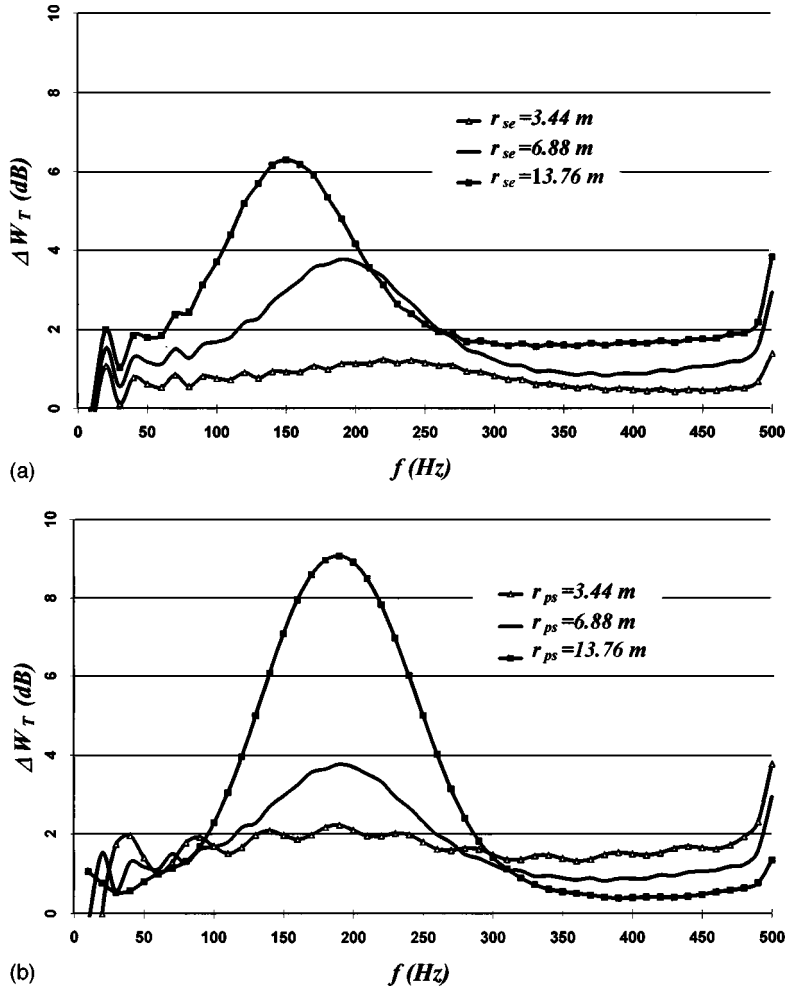


FIG. 6. The effective-frequency band for the control system with 15 control sources and error microphones, when (a) $r_{ps}=6.88$ m, $r_{se}=3.44$, 6.88, and 13.76 m, and (b) $r_{se}=6.88$ m, $r_{ps}=3.44$, 6.88, and 13.76 m.

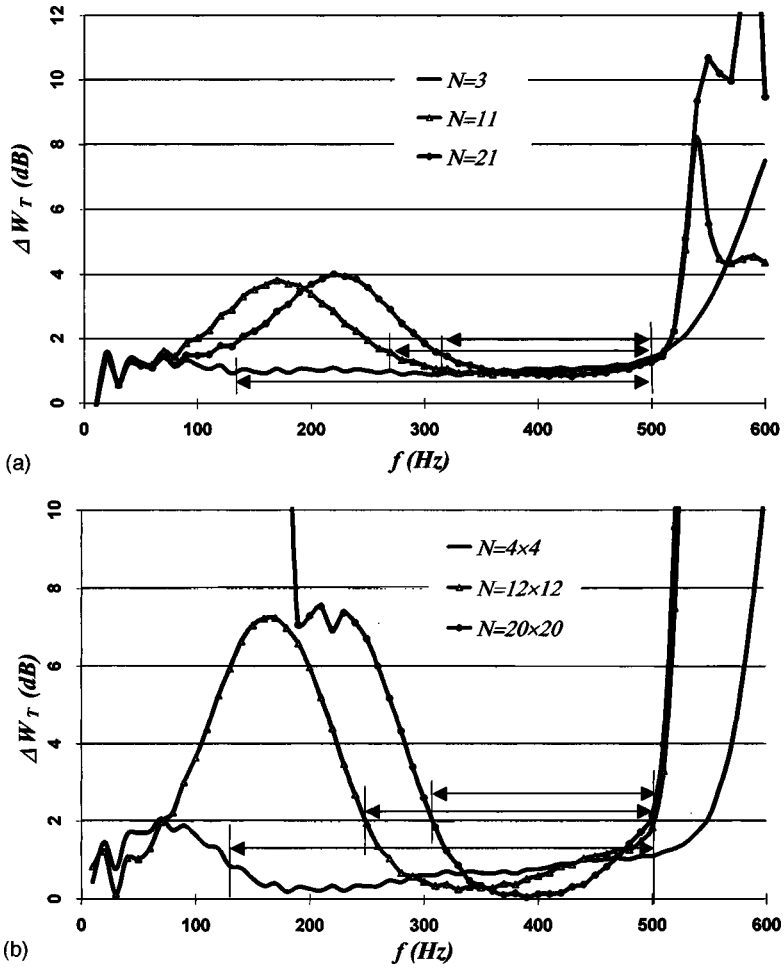


FIG. 7. Illustration of effective-frequency bands. (a) control system with sensors and actuators in two parallel lines, (b) control system in two parallel planes.

between the upper and lower limits, such as $r_{ss}=0.70$ m, the least increase of total power output of the control system is in the frequencies ranging from 320 to 720 Hz.

It is obvious that the width of effective-frequency band varies with the specific r_{ss} selected from the optimal spacing. They are, respectively, 720, 400, and 280 Hz for $r_{ss}=0.53$, 0.70, and 0.88 m in Fig. 5. Detailed observation of these three effective-frequency bands indicates that the ratios of upper and lower frequencies of the effective frequency band f_h/f_l for all the cases are about equal. The following task is to illustrate the relationship between this frequency ratio and the system configuration.

It has been found that the distances between the primary source and control-source array r_{ps} and between the control-source array and error-microphone array r_{se} have little effect on the effective-frequency width. Figure 6 shows an example to support this conclusion. In Fig. 6, the control system has 15 control sources and 15 error microphones, arranged as shown in Fig. 1. The distance between the primary source and control-source array r_{ps} is fixed as 6.88 m, and three

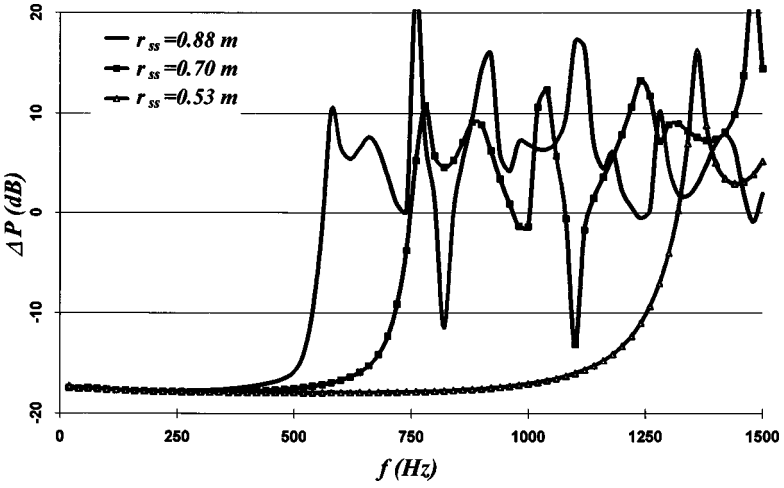


FIG. 8. Sound-pressure attenuation with respect to the frequency at a position behind the error microphones for three spacings of the control sources.

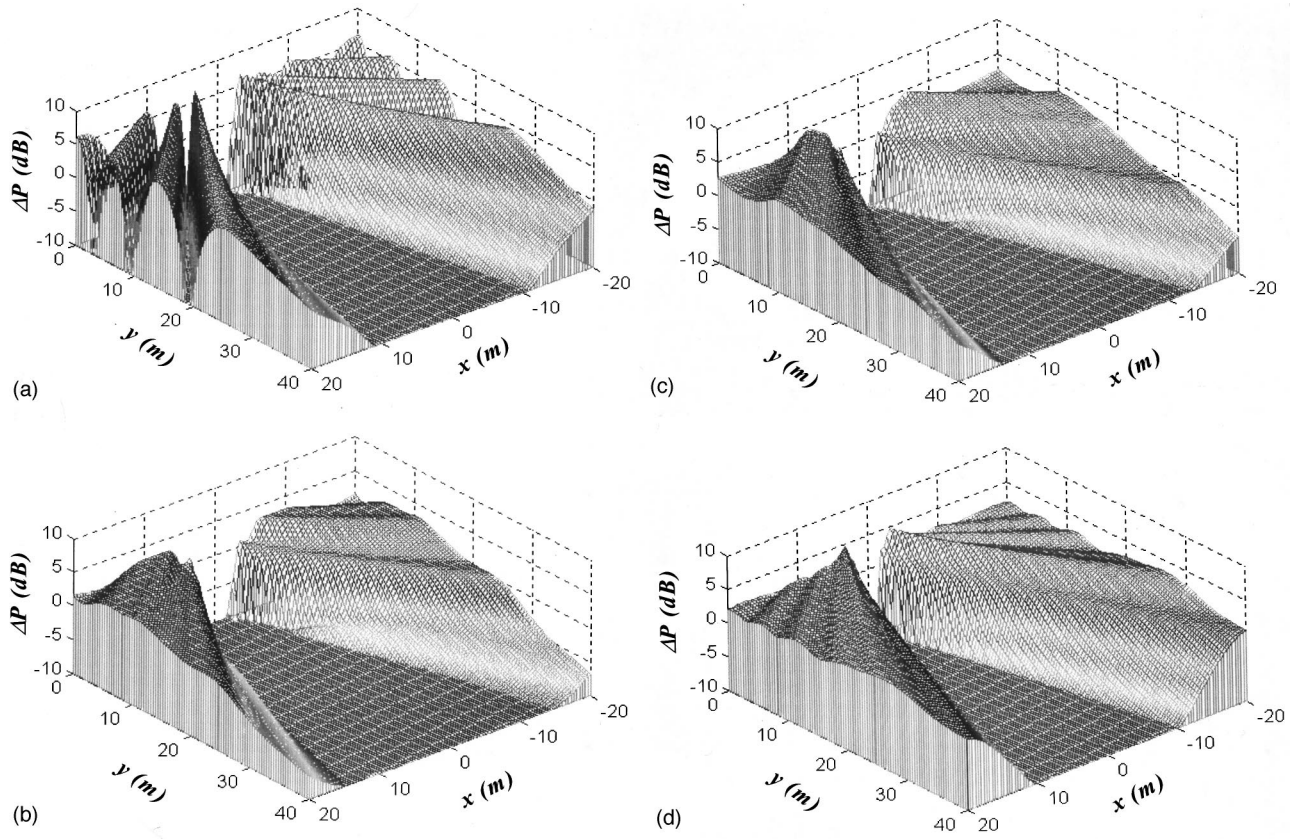


FIG. 9. Sound-pressure attenuation in an $x-y$ plane behind the error-microphone array for the control system with nine control sources and nine error microphones equally spaced in two lines when the noise is (a) pure tone with $f=500$ Hz and $r_{ss}=0.88$ m, (b) broadband with frequency range (220–500 Hz) and $r_{ss}=0.88$ m, (c) broadband with frequency range (320–720 Hz) and $r_{ss}=0.70$ m and (d) broadband with frequency range (500–1220 Hz) and $r_{ss}=0.62$ m.

distances of r_{se}, respectively as 3.44, 6.88, and 13.76 m, are chosen in Fig. 6(a). In Fig. 6(b), r_{se} is fixed as 6.88 m, and r_{ps} is chosen as 3.44, 6.88, and 13.76 m, respectively. The effective-frequency width is calculated at the upper limit of the optimal spacing range for the frequency 500 Hz.

It is clear that r_{se} and r_{ps} have little effect on the width of the effective-frequency band for this control system. The width of the effective-frequency band stays about the same for different r_{se} and r_{ps}. Comparing Fig. 6 to Fig. 5, the width of the effective-frequency band decreases. It seems that the number of control sources and error microphones affects the width of the effective-frequency band. Figure 7 shows how the width of the effective-frequency band decreases with the increase of control channels for a control system. The control system is arranged in two parallel lines in Fig. 7(a) and in two parallel planes in Fig. 7(b), with the distances between the sources and error microphones as $r_{ps}=6.88$ m, and $r_{se}=6.88$ m. The upper limit of optimal spacing range for 500 Hz defined by Eq. (5) is chosen as the spacing for control sources and error microphones. The width of the effective-frequency band of the control system is calculated for three different control channels, $N=3$, $N=11$, and $N=21$ in Fig. 7(a) and for three different control channels, $N=4\times4$, $N=12\times12$, and $N=20\times20$ in Fig. 7(b). The decrease of the effective-frequency bandwidth with the increase of the control channels is clearly shown in Fig. 7 for both control systems with control sources and error microphones equally spaced in two parallel lines and in two parallel planes.

Numerous numerical simulations for various configurations and numbers of the control sensors and actuators have enhanced this conclusion. It has been found that the width of the effective frequency range is determined not only by the number of control sources and error microphones, but also by if the number is odd or even. The upper and lower limits of the effective-frequency band have been calculated using the same relative level derivation $[\Delta=r_{se}/(r_{se}+r_{ps})]$ as described in Ref. 2. Then, the frequency ratio of the effective-frequency band for both systems shown in Fig. 1 and Fig. 2 can be fitted and given as

$$r_f=f_h/f_l=\begin{cases}\sqrt{1+\dfrac{36}{N}}, & N=3,5,7,...,\\ \sqrt{1+\dfrac{36(N+1)}{N(N-1)}}, & N=2,4,6,... .\end{cases} \tag{8}$$

Therefore, when the upper or lower frequency of the effective-frequency band is known, the width of the effective-frequency band can be calculated by

$$\Delta f=f_h(r_f-1)/r_f, \tag{9}$$

or

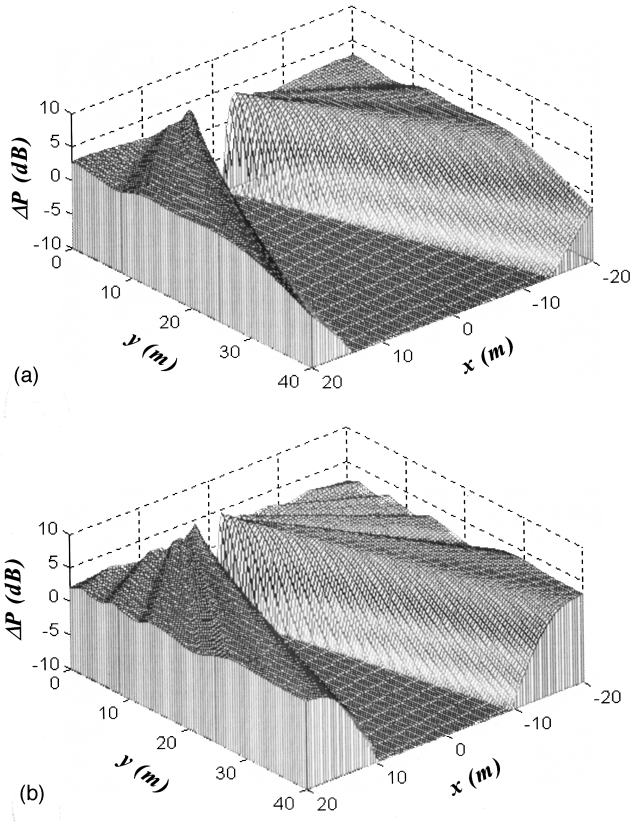


FIG. 10. Sound-pressure attenuation in an $x-y$ plane behind the error-microphone grid for the control system with 5×5 control sources and 5×5 error microphones equally spaced in two planes when the noise is (a) broadband with frequency range (220–500 Hz) and $r_{ss}=1.24$ m, and (b) broadband with frequency range (500–1220 Hz) and $r_{ss}=0.83$ m.

$$\Delta f = f_l(r_f - 1). \tag{10}$$

The effective-frequency band shown above is mainly from the analysis of total power increase of the control system. However, the effective-frequency band needs to be further verified in terms of creating the large area of quiet zone.

C. Zone of quiet

The ability to create a large area of quiet zone is one of the major indicators of a good local control system. The research on the multichannel control system for pure-tone noise has found that the configurations creating the large area of quiet zone are also those having the low increase of total power output. The sound-pressure attenuation caused by the control system in the space can be calculated by

$$\Delta P = 20\log_{10}(|P_T|/|P_p|), \tag{11}$$

where P_T is the resultant sound pressure at any position of the space expressed by Eq. (2), and P_p is the sound pressure caused by the primary source when the control system is off, which can be expressed by $P_p = Z_p q_p$. The previously discussed control system with nine control sources and nine error microphones equally spaced in two parallel lines is used as an example in Fig. 8, with the sound-pressure attenuation at a position in the central line of the control-source array and error-microphone array, and 20-m away from the error-microphone array for frequencies ranging from 0–1500 Hz.

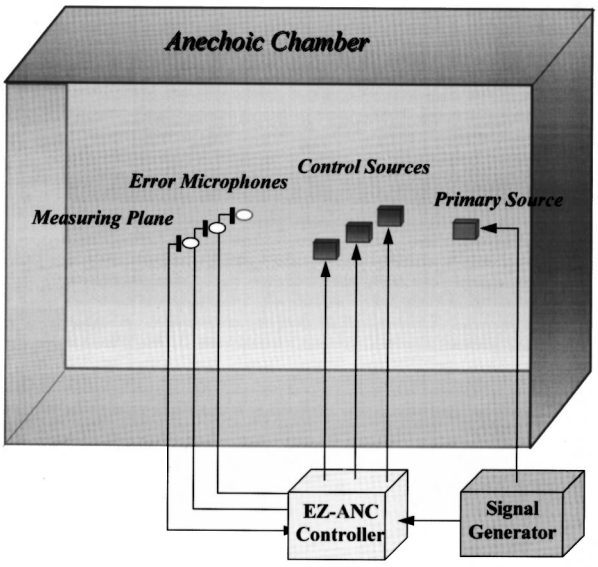


FIG. 11. Experimental setup.

It can be seen that the maximum attenuation of the sound pressure at that position can be as big as about 20 dB. When the upper limit of the control-source spacing $r_{ss}=0.88$ m is chosen, large sound attenuation can be achieved for frequencies lower than 500 Hz. For frequencies higher than 500 Hz, many frequency components of the sound pressure even increase. This increase can be as high as 20 dB or more. When the control spacing increases to the middle of the optimal range as $r_{ss}=0.70$ m, and to the lower limit of the optimal range as $r_{ss}=0.62$ m, the higher end of the effective frequency increases to 750 and 1250 Hz, respectively. However, the results shown in Fig. 8 do not mean that the effective-frequency band covers the whole range lower than 750 and 1250 Hz, respectively. This is because the total power increase can be large for low frequencies, as shown in Fig. 5; the sound pressure may have a big increase in other places of the space.

The sound-pressure attenuations for the noise of the effective-frequency band in an $x-y$ plane immediately behind the error-microphone array are demonstrated in Fig. 9. Figure 9(a) is the quiet zone created by the control system with $r_{ss}=0.88$ m, the upper limit of the optimal spacing for 500 Hz, and the noise is a pure-tone noise with frequency $f=500$ Hz. In Fig. 9(b), the spacing is $r_{ss}=0.88$ m, but the noise is of the wideband frequency ranging from 220 to 500 Hz. Figure 9(c) and (d) are the cases where $r_{ss}=0.70$ m, frequency band (320–720 Hz) and $r_{ss}=0.62$ m, frequency band (500–1220 Hz), respectively. The frequency bands used in these three cases [Fig. 9(b), (c), and (d)] are the effective-frequency band defined by the previous analysis of total power output [Eq. (8)].

The ability in creating a large quiet zone for the noise in the effective-frequency band is also demonstrated for a control system with control sources and error microphones equally spaced in two parallel planes. The control system presented here has 5×5 control sources and 5×5 error mi-

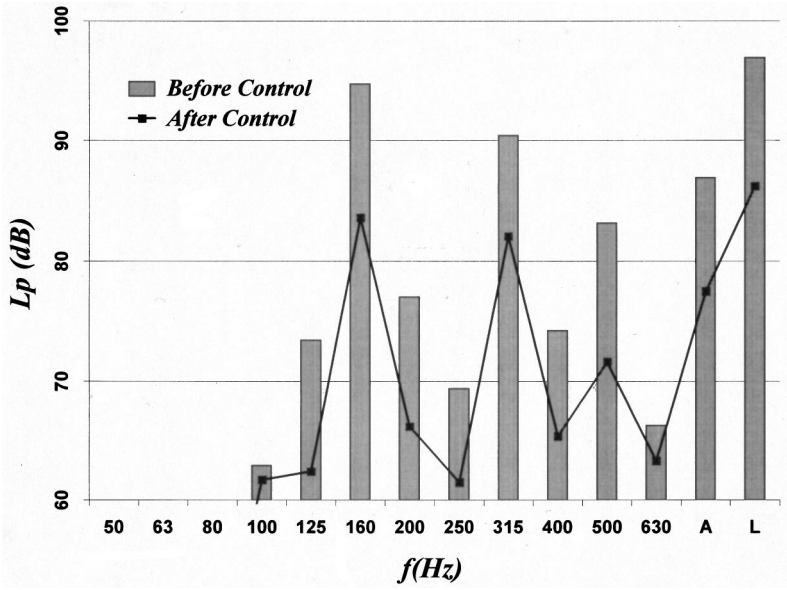


FIG. 12. The 1/3-octave spectra of sound attenuation for broadband noise.

crophones. The distance from the primary source to the control-source grid is $r_{ps}=3.44$ m, and the distance between the control-source grid and error-microphone grid $r_{se}=6.88$ m. It can be obtained from Eqs. (5) and (7) that the upper and lower limits of optimal spacing for 500 Hz are $r_{ss\text{-max}}=1.24$ m and $r_{ss\text{-min}}=0.83$ m, respectively. The corresponding effective-frequency bands in terms of upper and lower limits are given by Eq. (8) as (170–500 Hz) and (500–1430 Hz). The sound-pressure attenuation for the noise in these two effective-frequency bands is calculated in an $x-y$ plane immediately behind the error-microphone grid, and is shown in Fig. 10. For the case illustrated in Fig. 10, the system is arranged at its upper and lower limits of optimal spacing for 500 Hz, respectively.

It is obvious that the control system is efficient in creating a quiet zone for a broadband noise if the band is within the effective-frequency range. Compared with the quiet zone created for pure-tone noise, it can easily be seen that the size of the quiet zone created for broadband noise is comparable or even much larger than that for pure-tone noise. The sound-pressure fluctuation also becomes smaller in the area outside the quite zone.

III. EXPERIMENTAL VERIFICATION

Experiments were carried out in an anechoic chamber $4.2\times4.2\times4.2$ m^3 in size. Due to the limitation of the controller in terms of the number of channels available, a simple control system with three control sources and three error microphones equally spaced in two parallel lines was tested. The primary noise source was a half-enclosed loudspeaker. Three half-enclosed speakers and three microphones were used as control sources and error sensors, respectively. A multichannel EZ-ANC, which is adaptive with the filter-x LMS method, was applied as the controller.[9] The signal generated by the signal generator was fed directly into the primary source, and was also provided to the controller as a reference signal. Three control channels of the controller were used to cancel the total sound pressure at the positions of three error microphones. Parameters of the controller were: sampling frequency $f_s=5210$ Hz; step size $\mu=0.01$ (fixed); number of weights=32; leakage factor=0.0001. The weights were updated every sample.

The arrangement of the control system is shown in Fig. 11, where the three control sources and the three error microphones were placed in two parallel lines. The distance from the primary source to the control-source array is $r_{ps}=0.688$ m, and the distance from the control-source array to the error-microphone array is $r_{se}=1.376$ m. The position of the primary source is at $[-(r_{ps}+r_{se}),\ 0,0]$ and the coordinates of the control sources and error microphones are $[-r_{se},(i-2)r_{ss}/2,0]$ and $[0,(i-2)r_{ss}/2,0]$, respectively, where $i=1,2,3$.

For the noise frequency of $f=500$ Hz, according to Eqs. (5) and (6), the optimal spacing range of the system is $0.23\ m\leqslant r_{ss}\leqslant0.87$ m. The spacing of the control sources and error microphones is chosen within the optimal range as $r_s=0.85$ m, which is about the upper limit of the optimal range. The effective-frequency band for this configuration can be calculated by Eq. (8) as (160–500 Hz). The measurement of sound-pressure attenuations was carried out in an $x-y$ plane behind the control-source array. Sound-pressure levels at 11×10 positions in the measuring plane were measured before and after the control. To demonstrate that the optimally arranged multichannel control system is also effective to broadband noise, a pure-tone noise of 500 Hz and a broadband noise with frequency range of 160–500 Hz were tested for the comparison. Because of the limitation of the controller and number of filter coefficients available for each channel, the broadband noise is replaced by a ramp noise with the fundamental frequency of 160 Hz; a low-pass filter with cutoff frequency of 600 Hz is used to get rid of the high-frequency components. There are three components in the frequency range (160–500 Hz). Figure 12 shows the 1/3-octave spectra of sound attenuation at a position in the measuring plane ($0.6m$, 0, 0) for broadband noise. It is clear that the sound-pressure attenuation is obtained in each frequency component, which varies from 2 to more than 10 dB. The

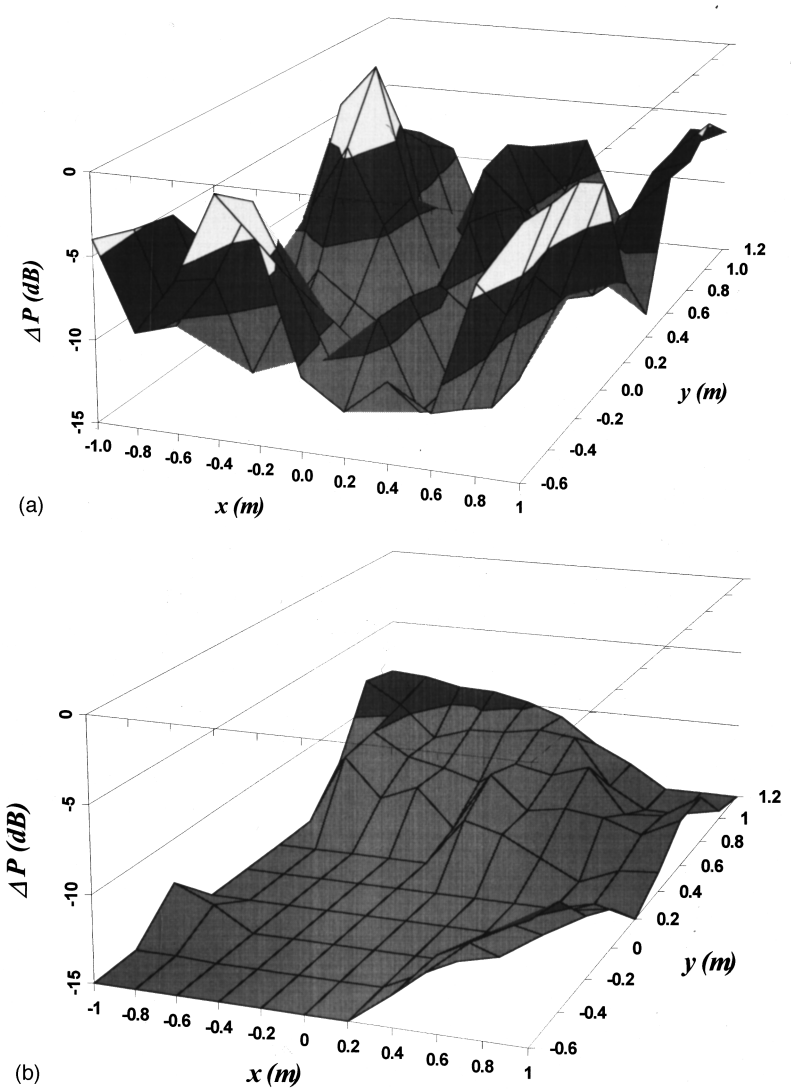


FIG. 13. Experimental results of sound-pressure attenuation in an $x-y$ plane behind the control sources when the noise is (a) pure tone with $f=500$ Hz, and (b) ramp noise with frequency range (160–500 Hz).

total sound-pressure attenuation is mainly determined by the attenuation of peak components of the noise.

The sound-pressure attenuations in the measuring plane for both pure-tone noise and broadband noise are given in Fig. 13. Large sound-pressure attenuations were achieved in both cases in the measuring plane with the size 2×1.8 m^2. Although the attenuation for the pure-tone noise is more than 10 dB in most of the area, there are some areas around the corner of the plane where the sound attenuation can be less than 5 dB. To the broadband noise, the sound-pressure attenuation is more than 10 dB over almost all of the area; it even exceeds 15 dB over most of the area.

The ability of a control system in creating a large quiet zone for a noise with multiple tone within a frequency range (ramp noise with fundamental frequency at 160 Hz for this case) was experimentally demonstrated. The size of the quiet zone is even larger than that of the pure-tone noise. This has also been verified in research on the application of an active control system to improve the insertion loss of the noise barrier for a band of low-frequency noise.[10] It should be noted that although an optimal arranged control system has been found effective for broadband noise, which means that the control system can attenuate all the frequency components covering the frequency band, this ability of the control system can be limited by the controllers in practical application.

IV. CONCLUSIONS AND DISCUSSIONS

An optimally arranged multichannel active noise-control system is effective not only for specific pure-tone noise, but also for noise with a frequency range. The effective range of the broadband noise is dependent on the geometrical spacing of the control sources, and on the number of control channels.

For the application of broadband-noise control in open space, the selection of optimal spacing and determination of effective bandwidth can follow this simple procedure: (1) determine the cutoff frequency of the band or the frequency of the highest peak f_h, (2) calculate the upper limit of the optimal spacing with respect to f_h, (3) choose the upper limit as the spacing of control sources and error microphones for the control system, (4) calculate the width of the effective-frequency band, (5) design a low-pass filter to get rid of frequency components higher than f_h (sometimes, a high-pass filter may also be necessary to get rid of too-low

frequency components), then the system will be capable of attenuating the noise and creating a large area of quiet zone for that frequency band.

Although the results presented in this paper show that the control system is efficient for broadband noise, in practical applications this efficiency can be undermined by the limitation of the controller. Further development of the multichannel digital controller for attenuating broadband noise may overcome this difficulty.

[1] J. Guo and J Pan, ''Analysis of active noise control in a free field,'' *Proc. Active 95*, Newport Beach, California, pp. 649–660 (1995).

[2] J. Guo, J. Pan, and C. Bao, ''Actively created quiet zones by multiple control sources in free space,'' J. Acoust. Soc. Am. **101**, 1492–1501 (1997).

[3] J. Guo and J. Pan, ''Further investigation on actively created quiet zones by multiple control sources in free space,'' J. Acoust. Soc. Am. **102**, 3050–3053 (1997).

[4] C. R. Fuller, S. J. Elliott, and P. A. Nelson, *Active Control of Vibration* (Academic, New York, 1996).

[5] A. David and S. J. Elliott, ''Numerical studies of actively generated quite zones,'' Appl. Acoust. **41**, 63–79 (1994).

[6] P. Joseph, S. J. Elliott, and P. A. Nelson, ''Near field zones of quiet,'' J. Sound Vib. **172**, 605–627 (1994).

[7] P. A. Nelson, A. R. D. Curtis, S. J. Elliott, and A. J. Bullmore, ''The minimum power output of free field point sources and the active control of sound,'' J. Sound Vib. **116**, 397–414 (1987).

[8] A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (The Acoustical Society of America, Woodbury, NY, 1989).

[9] S. D. Snyder and G. Vokalek, *EZ-ANC User's Guide* (Causal System Pty. Ltd., South Australia, 1994).

[10] J. Guo and J. Pan, ''Increasing the insertion loss of noise barriers using active control system,'' J. Acoust. Soc. Am. **104**, 3408–3416 (1998).

Measurement of transient response of rooms and comparison with geometrical acoustic models

J. S. Suh[a] and P. A. Nelson
Institute of Sound and Vibration Research, The University of Southampton, Southampton SO17 1BJ, United Kingdom

(Received 18 July 1997; revised 17 November 1998; accepted 7 January 1999)

The transient responses of rooms obtained from measurements and geometrical room acoustic models are compared. The conventional image model was modified in order to include interference in the modeling process by accounting for the amplitude and phase changes produced at each reflection. This ''phase image model'' has the capability to find the simulated pressure impulse response of a modeled room. In order to validate this model, some theoretical analyses and measurements were first carried out on a single reflecting surface. Finally the pressure impulse responses and energy impulse responses were measured in real rooms and these were compared with those from the phase image model and the conventional models. It is shown that the inclusion of complex reflections in the phase image model greatly increases the accuracy of the predictions. In this work Gabor pulses are used as the optimal source strength signals, enabling comparison between measurements and room acoustic models. © *1999 Acoustical Society of America.* [S0001-4966(99)02804-0]

PACS numbers: 43.55.Br, 43.55.Ka, 43.55.Gx [JDQ]

INTRODUCTION

The acoustical impulse response functions of a room can provide very useful information for the prediction of a room's acoustic response due to any arbitrary input characteristics, as well as for the control of its sound field. This implies that by convolving the reverberation-free sound signals with an impulse response function of a room which is to be built in the future, the acoustical impression of a room can be evaluated before its completion.

Currently two models, the ray tracing model[1–3] and the image model,[4–7] are being widely used for the modeling of the sound field in rooms. The early part of the impulse response, which is dominated by specular reflections, is modeled by an image model, and the later part of it, which is dominated by diffuse reflections, is modeled by a ray tracing model.[8,9] Considerable effort is also being made to model the scattered field.[10–13] Of course finite element methods and boundary element methods can be used, but their usage in practice is limited to low frequency regions. In fact, as far as the specular reflections are concerned, the conventional ray tracing and image models will produce identical energy impulse responses in the limiting case when the ray tracing model generates a large number of sound particles and at the same time uses a very small detector volume. In this paper these two models will be referred to as the conventional model.

Hitherto many room acoustic models have been developed to describe the various aspects of room acoustic fields. However, these models are not able to take full account of the inherent nature of sound waves to interfere, scatter, and diffract as discussed recently by Mortessagne and Legrand.[14] As far as the physics of the process is concerned, interference for instance, should be modeled by complex reflections. However, it has been reported that adding individual phases to the reflection factors hardly changes anything audible to the overall impulse response, and this point will be discussed further in Sec. IV.[15] Likewise particle diffusion modeling methods may not accurately describe the phenomena of complex scattering and diffraction, but several successful applications of these methods have been reported when room responses were averaged over a sufficient number of reflections.[16] One approach using the conventional ray tracing or image model is to compute the energy impulse response function associated with a given frequency band by simply adding the energies arriving as a function of time. The contribution from each reflection can be computed by simply multiplying the strength of the relevant image source by the band averaged energy reflection coefficient.

In this work an attempt has been made to take interference into account. At each reflection of a sound wave from a wall surface, a complex reflection coefficient was used to account for the amplitude and phase change. Also, the amplitude and phase change of a sound wave as it travels from a source position to a detector position are of course included. In the main text this model will be referred to as the ''phase image model.'' It is shown that this type of model gives a much better representation of the energy impulse response in a given frequency band than the conventional model with simple energy addition. Geest and Pätzold have also recently developed a new model including interference.[17] They compared the simulated room transfer functions computed with their model and a boundary element method in the frequency domain. Good agreement was found in a frequency range of 200 Hz–550 Hz for a parallelepiped room sized $2.5\times1.5\times1.2$ m.

An important issue to be addressed in modeling the

[a]Present address: Research & Development Center, Pyung Hwa Industrial Co., Ltd., 29-17, Bonri-Ri, Nongong-Eup, Dalsung-Gun, Taegu, Korea

complex reflections from a room surface is, of course, the correct choice of reflection coefficient. In this paper we begin from the exact solution for the reflection of waves radiated by a spherical source above an infinite plane boundary of finite impedance. The exact solution is expressed in terms of an inverse wave number transform and this is evaluated numerically in a number of representative cases. The results are then compared with the predictions of a simple ''image model'' of the reflected field which assumes that the reflected wave is generated by a spherically radiating image source. The strength of this image source is multiplied by the complex reflection coefficient associated with plane waves incident from the angle subtended at the surface by the original source. It is shown that the simple image model gives a good representation of the reflected field at positions that are at least several wavelengths from the source and provided that the angle of the incident field is not near grazing.

One of the ultimate goals of room acoustic modeling is to predict at least the major features of the impulse responses of existing rooms or rooms to be built in the future. In this respect, it is important to compare the simulated impulse responses from models with those from measurements in the time domain. Therefore it is necessary to define a method for comparing measurements with prediction from computer models since, in practice, it is very difficult to measure the true ''impulse response'' of an acoustic space. This is because it is very difficult to find an acoustic source of known strength which is both omnidirectional and which operates over a sufficiently broad frequency range to enable the generation of a repeatable transient output. In this work we make use of omnidirectionally radiating sources whose strength is easily measured and whose output is of limited bandwidth, but we make use of a particular form of transient source strength signal. The desirable characteristic of this signal is such that the time width should be the shortest possible for a given frequency bandwidth. Then the direct wave and each reflected wave can be better resolved. Since real acoustic sources become directional as frequency goes up, a frequency range showing good omnidirectionality for a particular source has to be chosen for the measurement of transient responses. The bandwidth of a source strength signal should be defined within this frequency range. It is demonstrated that the Gabor pulse[18,19] (a short duration pulse whose time envelope and spectrum are Gaussian functions of time and frequency, respectively) can best meet these requirements.

I. SINGLE REFLECTING SURFACE PROBLEM

In this section, the validity of using the phase image model is investigated both analytically and experimentally in the simplest acoustic environment, i.e., in the half-space, as shown in Fig. 1.

A. Comparison between the analytical solution and the phase image model

The complex acoustic pressure $p(x,y,z)$ at the detector position in Fig. 1, when the source strength (volume acceleration) is unity, is given by

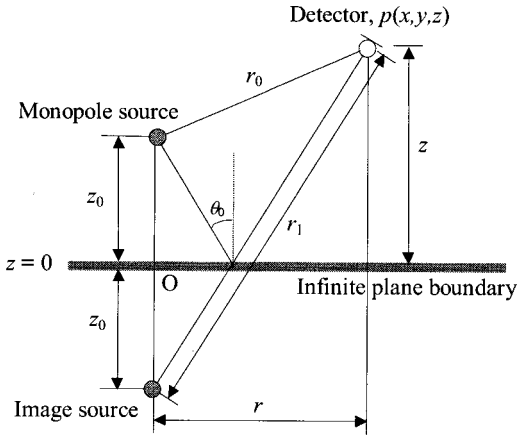


FIG. 1. Configuration of source and detector in a half space. O is the origin of the coordinate system.

$$p(x,y,z)=\frac{\rho_0 e^{-jkr_0}}{4\pi r_0}+p_{\text{refl}}, \tag{1}$$

where p_{refl} is the reflected wave field, k is the wave number of the spherical wave radiated from the monopole source, and ρ_0 is the ambient density of air. By decomposing a spherical wave into an infinite number of plane waves p_{refl} can be expressed by[20]

$$p_{\text{refl}}=-\frac{j\rho_0 k}{4\pi}\int_{\Gamma_\theta} J_0(kr\sin\theta)e^{-jk(z+z_0)\cos\theta}R(\theta)\sin\theta\, d\theta, \tag{2}$$

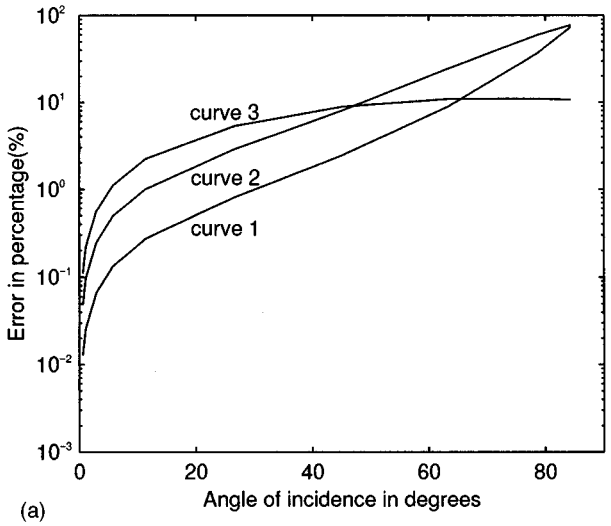
where $R(\theta)$ is the complex reflection coefficient of the infinite plane boundary which depends upon the angle of incidence, θ, and $J_0(kr\sin\theta)$ is the Bessel function of zero order. Then Eq. (2) can be computed by numerical integration.[21,22] It can be easily demonstrated that, except near-grazing incidence, Eq. (2) converges rapidly. The numerical technique used in this work is the Trapezoidal rule. The magnitude of the constant step $\Delta\theta$ replacing $d\theta$ in Eq. (2) should be made as small as possible for good accuracy. A value of 10^{-5} radians was chosen for $\Delta\theta$ in this work. In the Appendix, the underlying principles of the numerical integration used in this paper are presented and a specific result is shown which demonstrates the convergence of the integral.

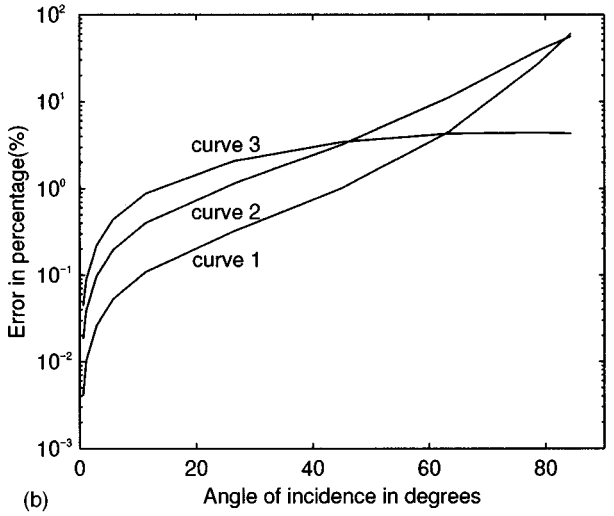
The reflection coefficient $R(\theta)$ of the infinite plane boundary, when a plane wave is incident upon it, can be expressed by

$$R(\theta)=\frac{\xi\cos\theta-1}{\xi\cos\theta+1}, \tag{3}$$

where ξ is the specific acoustic impedance of the infinite plane boundary normalized by the characteristic impedance $\rho_0 c$ of air, and c is the speed of sound. In this paper, it is assumed that all the acoustic materials are locally reacting. In most cases, except the near-grazing incidence, p_{refl} can be replaced with good approximation by

$$p_{\text{image}}=\frac{\rho_0 e^{-jkr_1}}{4\pi r_1}R(\theta_0). \tag{4}$$





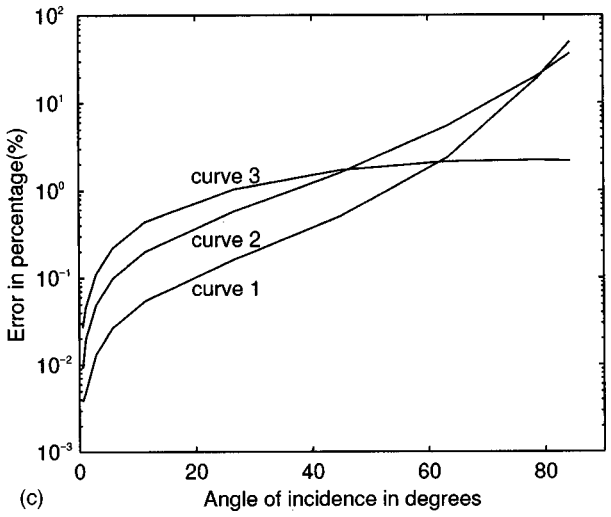


FIG. 2. (a) The percentage (%) error between p_{refl} and p_{image}. The angle of incidence denotes θ_0 in Fig. 1. r is 2λ. (b) The percentage (%) error between p_{refl} and p_{image}. The angle of incidence denotes θ_0 in Fig. 1. r is 5λ. (c) The percentage (%) error between p_{refl} and p_{image}. The angle of incidence denotes θ_0 in Fig. 1. r is 10λ.

This is the contribution to the total field made by the image source as illustrated in Fig. 1. The percentage error (%) between p_{refl} and p_{image} (the error due to the plane wave approximation) can be defined as

TABLE I. The acoustic properties of the infinite plane boundary.

	ξ	α_d
curve 1	$5.00-j11.00$	0.21
curve 2	$1.00-j2.83$	0.41
curve 3	$0.59+j0.57$	0.71

$$\frac{|p_{\text{refl}}-p_{\text{image}}|}{|p_{\text{refl}}|}\times 100(\%).$$

In Fig. 2(a), (b), and (c), this is plotted against the angle of incidence θ_0 for three different values of ξ when r is 2λ, 5λ, 10λ, respectively where λ is the wavelength. The ξ's are tabulated in Table I together with the corresponding diffuse absorption coefficient, α_d, when local reaction is assumed.[23] It should be noted that the results illustrated in Fig. 2(a), (b), and (c) are independent of frequency since the values of r, r_1, z_0, and z were given in terms of λ.

The results demonstrate that the degree of accuracy achieved by using Eq. (4) to define the reflected field is most sensitive to the angle of incidence θ_0. When θ_0 is smaller than 60°, the true reflected field can be replaced by Eq. (4) with good accuracy provided that the detector position is many wavelengths from the source. In general, we reach the important conclusion that Eq. (4) is a good approximation to the reflected field except for grazing incidence sound or surfaces of high absorption coefficient when the source–receiver distance is less than or equal to a few wavelengths.

B. Comparison of the acoustical frequency response functions between measurement and the phase image model

The measurement setup is drawn in Fig. 3. A 5-cm-thick plastic foam, which is an open cell type, is used to cover a 5-cm-thick high-density fiber board. The intention was to simulate a rigid-backing condition. The monopole source used in the measurement is shown in Fig. 4.

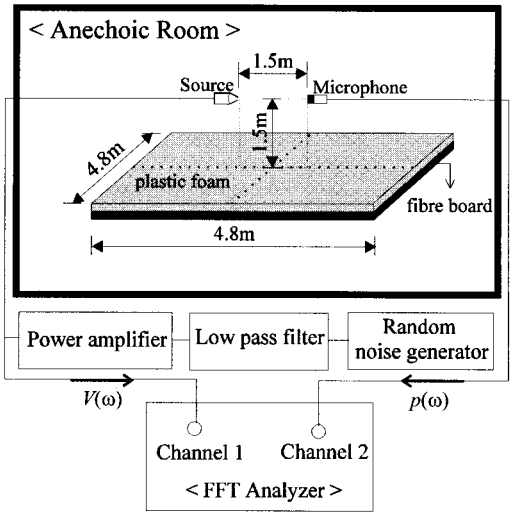


FIG. 3. Measurement setup, $p(\omega)$ is the pressure at microphone position and $V(\omega)$ is the output drive voltage of power amplifier.

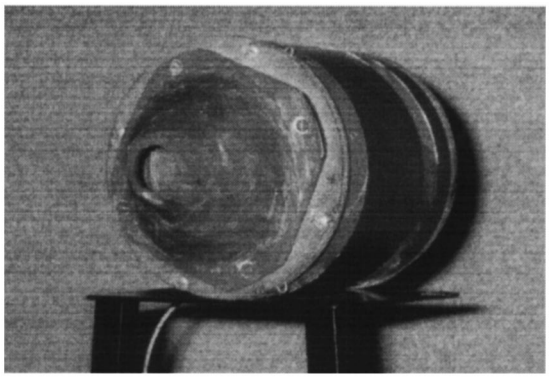


FIG. 4. Monopole source used in the measurement. The opening diameter of the horn is 28 mm. A loudspeaker unit 15 cm in diameter is fixed between the horn and the cylindrical body. (Courtesy of Dr. K. R. Holland and Professor F. J. Fahy, University of Southampton.)

Its directivity will depend upon the polar angle measured from the central axis. Let us denote the directivity of the source as $D(\omega,\theta°)$ when the angular frequency is ω and the polar angle is $\theta°$. As an illustration $D(\omega,63°)/D(\omega,0°)$ was measured in an anechoic room, and is plotted in Fig. 5. This source shows good omnidirectionality over the whole range of polar angles on the frequency range of 80 Hz–800 Hz.

The frequency response function between the pressure and the output drive voltage of the power amplifier in Fig. 3, $p(\omega)/V(\omega)$, is plotted in Fig. 6. In order to obtain the acoustical frequency response function which is the pressure response per unit source strength, a source calibration was carried out in an anechoic chamber using the same measurement setup that is shown in Fig. 3. In this paper, volume acceleration is chosen to be the source strength. Like the previous measurement, $p(\omega)/V(\omega)$ was measured. This is shown in Fig. 7. The volume acceleration per unit output drive voltage of the power amplifier, $\dot{q}(\omega)/V(\omega)$, can be computed from

$$\frac{\dot{q}(\omega)}{V(\omega)}=\frac{4\pi r_{\mathrm{cal}}\exp(jkr_{\mathrm{cal}})}{\rho_0}\cdot\frac{p(\omega)}{V(\omega)}, \tag{5}$$

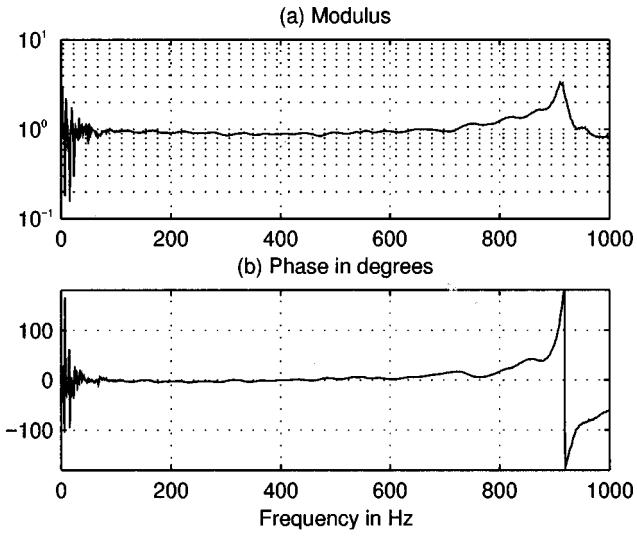


FIG. 5. Relative directivity $D(\omega,63°)/D(\omega,0°)$ of the source shown in Fig. 4.

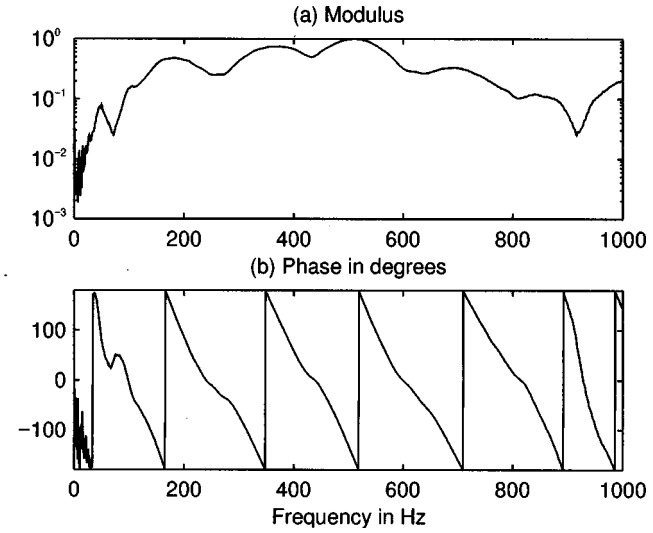


FIG. 6. The frequency response function between the pressure and the output drive voltage of the power amplifier for the single surface measurement. (Refer to Fig. 3.)

where $\dot{q}(\omega)$ denotes the time derivative of the volume velocity, and r_{cal} is the distance between the source and the microphone. In this measurement, r_{cal} was 2.3 m. In order to produce the simulated frequency response function from the phase image model, the normal specific impedance of the plastic foam, ξ, was measured by an impedance tube test with a rigid-backing condition. The results are plotted in Fig. 8. This enabled $R(\theta_0)$ to be calculated from Eq. (3). The acoustical frequency response function can be obtained by dividing the frequency response function shown in Fig. 6 by Eq. (5). This is shown in Fig. 9 together with that simulated by the phase image model. Excellent agreement is shown.

C. Comparison of the pressure impulse responses between measurement and phase image model

It is useful in dealing with room acoustic models to compare the time domain responses. The impulse response

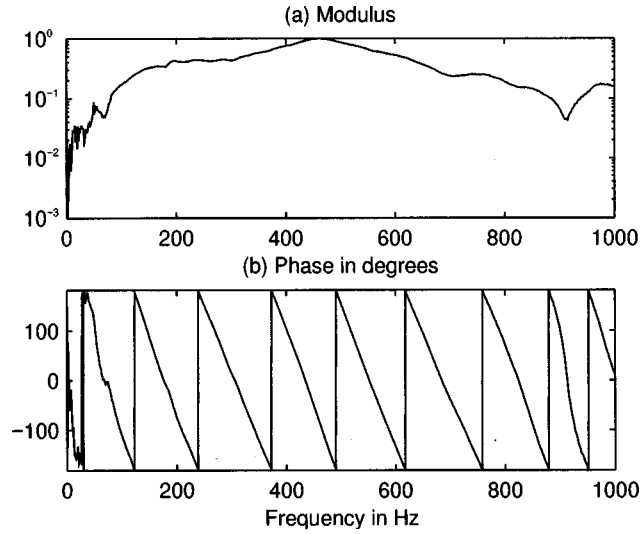


FIG. 7. The frequency response function between the pressure and the output drive voltage of the power amplifier in an anechoic chamber.

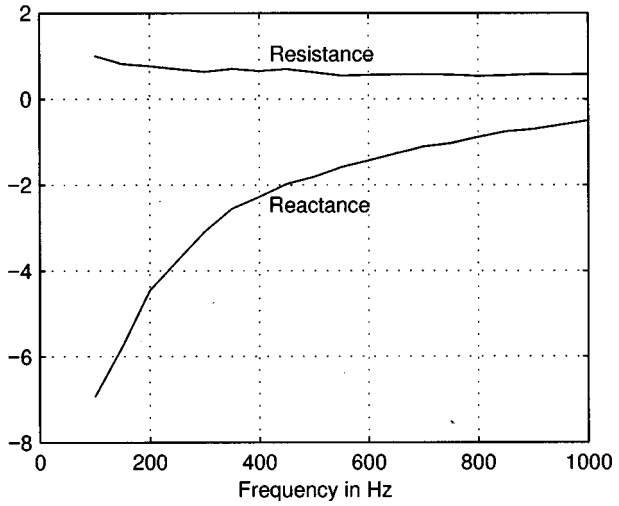


FIG. 8. Normal specific acoustic impedance of the plastic foam.

functions are defined for an omnidirectional source. Also, real acoustic sources are reasonably omnidirectional only within limited frequency bandwidth. Therefore a suitable source strength signal defined within this bandwidth should be used. Here, we illustrate the importance of the choice of the form of this limited bandwidth signal.

For any band limited time signal, the product of the time width Δt and its frequency bandwidth $\Delta\omega$ (or Δf) is always subject to the restriction of the ''Uncertainty Principle'' which states that

$$\Delta t \cdot \Delta\omega \geqslant \pi \quad \text{or} \quad \Delta t \cdot \Delta f \geqslant 0.5, \tag{6}$$

where ω is in radians/s and f is in cycles/s. We illustrate the importance of using a signal whose value of Δt is as small as possible for a given $\Delta\omega$ by presenting the results of using two different signals. The first source strength signal to be used is $\psi_1(t)$ given by[24]

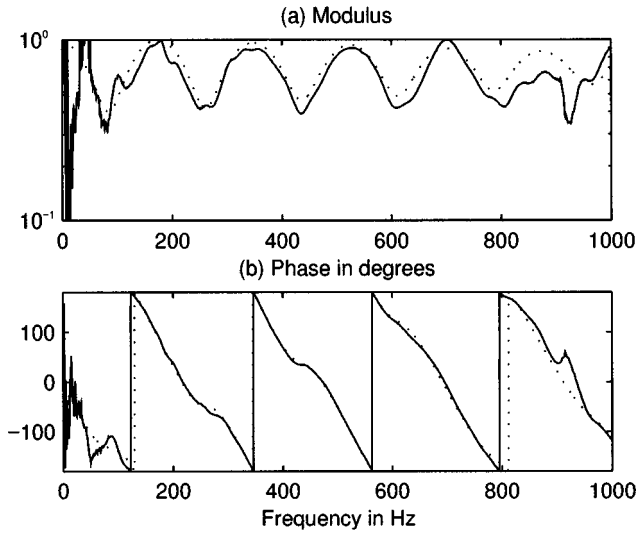


FIG. 9. The acoustical frequency response functions obtained from the measurement and the phase image model. Measurement: solid line; Phase image model: dotted line.

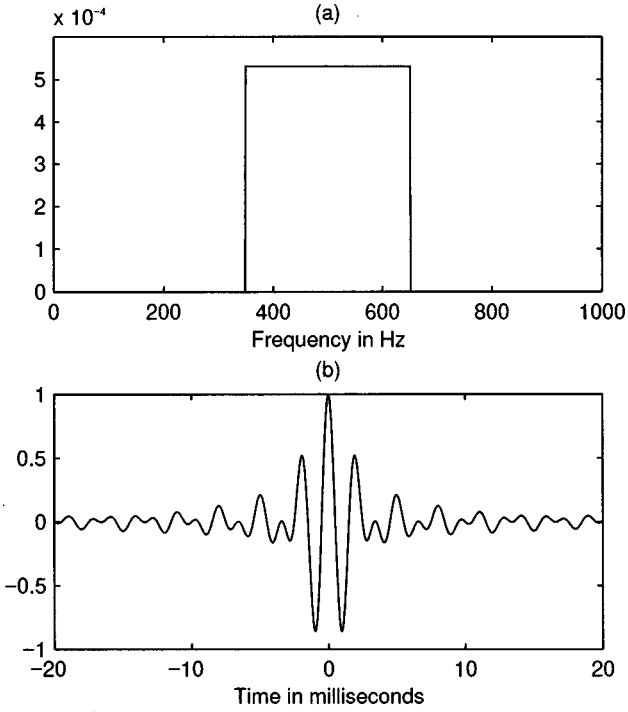


FIG. 10. Source strength signal. (a) $\phi_1(f)$, and (b) $\psi_1(t)$.

$$\psi_1(t) = \frac{\sin\frac{1}{2}(\Delta\omega)t}{\frac{1}{2}(\Delta\omega)t}\cos(\omega_0 t), \tag{7}$$

where the product $\Delta t \cdot \Delta\omega$ takes the value of 2π and ω_0 is the center frequency of the band. Thus Δt is twice the minimum value possible for the given value of $\Delta\omega$. Its Fourier transform $\phi_1(\omega)$ is a square spectrum given by

$$\phi_1(\omega) = \begin{cases} \dfrac{1}{\Delta\omega} & \text{for } \omega_0 - \dfrac{\Delta\omega}{2} \leqslant \omega \leqslant \omega_0 + \dfrac{\Delta\omega}{2} \\ 0 & \text{for } \omega_0 - \dfrac{\Delta\omega}{2} > \omega > \omega_0 + \dfrac{\Delta\omega}{2} \end{cases}. \tag{8}$$

The second source strength signal is a Gabor pulse, $\psi_2(t)$, given by

$$\psi_2(t) = e^{-\alpha^2 t^2}\cos(\omega_0 t). \tag{9}$$

$\psi_2(t)$ is the real part of the Gabor elementary signal,[18,19] $g(t)$, given by

$$g(t) = e^{-\alpha^2 t^2} e^{j\omega_0 t}, \tag{10}$$

where the product $\Delta t \cdot \Delta\omega$ takes the minimum value of π. For $g(t)$, the constant α is related to Δt, $\Delta\omega$ by

$$\Delta t = \sqrt{\frac{\pi}{2}}\frac{1}{\alpha}, \quad \Delta\omega = \sqrt{2\pi}\alpha. \tag{11}$$

For $\psi_2(t)$, however, the product $\Delta t \cdot \Delta\omega$ becomes equal to π. Its Fourier transform $\phi_2(\omega)$ is given by

$$\phi_2(\omega) = \frac{1}{2}\frac{\sqrt{\pi}}{\alpha}\left[e^{-\pi^2[(\omega+\omega_0)^2/4\alpha^2]} + e^{-\pi^2[(\omega-\omega_0)^2/4\alpha^2]}\right]. \tag{12}$$

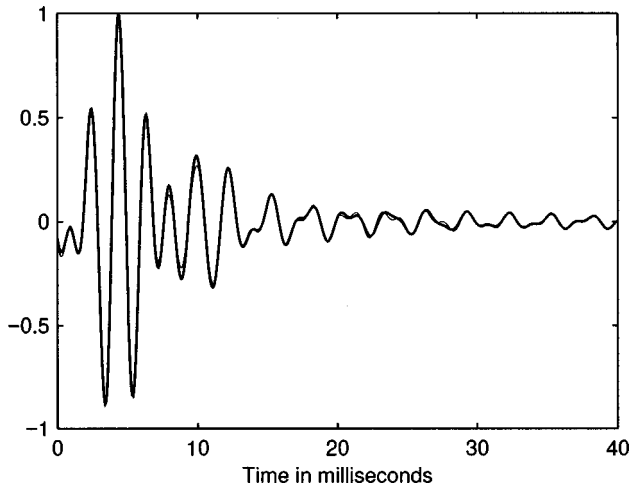


FIG. 11. Pressure impulse responses due to $\psi_1(t)$. Measurement: Thick solid line; Phase image model: Thin solid line.

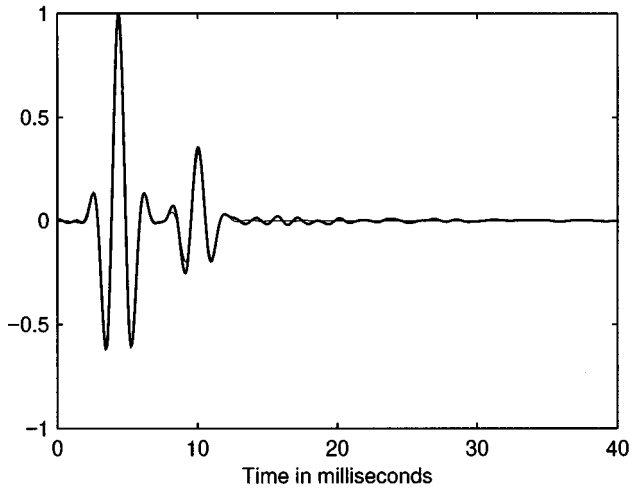


FIG. 13. Pressure impulse responses due to $\psi_2(t)$. Measurement: Thick solid line; Phase image model: Thin solid line.

For the single surface reflection experiment described above, Fig. 10 shows $\psi_1(t)$ and $\phi_1(f)$ with $\Delta t = 1/300$ s, $\Delta f = 300$ Hz, and $f_0 = 500$ Hz. The pressure impulse responses due to $\psi_1(t)$ are compared in Fig. 11. Figure 12 shows $\psi_2(t)$ and $\phi_2(f)$ with $\Delta t \cong 1/600$ s, $\Delta f \cong 300$ Hz, $f_0 = 500$ Hz. The pressure impulse responses due to $\psi_2(t)$ are compared in Fig. 13. These results were computed by taking the inverse Fourier transform of the product of the acoustical frequency response function and the Fourier transform of the source strength signal.

For both of the input signals the phase image model predicts the measured impulse responses with good accuracy even though the incident angle dependent reflection coefficient was deduced from a local reaction assumption. This will validate the use of the phase image model in predicting the early part of the impulse responses of rooms when the major dimensions of the wall surfaces are greater than the acoustic wavelength considered, i.e., at relatively high frequencies. A Gabor pulse has the advantage over other types of signal that its time duration is the shortest possible for a given frequency bandwidth. Figure 13 in particular shows how well the direct and first reflected pulses are resolved when the source strength time history is a Gabor pulse. As will be shown in the next section, a Gabor pulse is a very useful source strength signal for comparing the impulse responses of rooms obtained from measurements and models. The results shown in Fig. 11 give an example of the poor resolution in the time domain which would be produced by attempting the same type of analysis in, for example, octave bands.

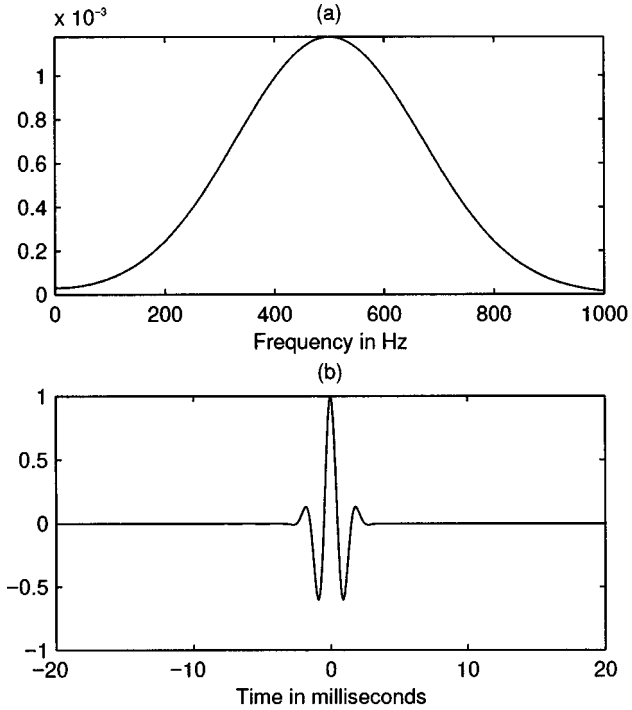


FIG. 12. Source strength signal. (a) $\phi_2(f)$, and (b) $\psi_2(t)$.

II. GEOMETRICAL ROOM ACOUSTIC MODELS

In this section, the basic principles of the conventional ray tracing and image models, and the phase image model are briefly explained. The method of computing the energy impulse response from the conventional ray tracing and image models is described. The modifications necessary to predict both the pressure and energy impulse responses from the phase image model are also presented. The computer programs for the three models were written in C programming language. These were run on a Silicon Graphics Indigo R4000 Unix workstation, whose performance is 86 MIPS, 16 MFLOPS, and whose memory space is 48 Mbytes.

A. Ray tracing model

A number of sound ray ''particles'' of infinitesimally small size are released from a monopole source at one time instant $t=0$. They are released omnidirectionally and uniformly. Their unit direction vectors can be computed from a set of equations presented by Krokstad *et al.*[1] The minimum number of the ray particles that should be released, when a sphere detector is used, can be calculated from[7]

$$N_{\min} = 4(ct_{\max})^2/r_k^2, \tag{13}$$

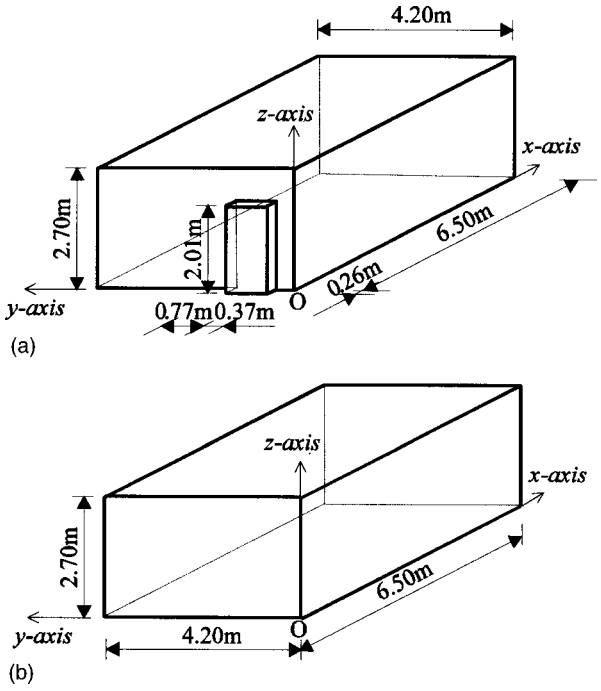


FIG. 14. (a) The actual configuration of medium-sized room. O is the origin of the coordinate system. (b) The simplified configuration of medium-sized room for modeling. O is the origin of the coordinate system.

where $t_{\max}$ is the required duration of the impulse response and r_k is the radius of the sphere detector. Each ray particle initially carries a certain amount of energy and moves in a straight line with the speed of sound between two successive reflections. When a ray particle hits a reflecting surface its energy is decreased by the diffuse absorption coefficient of the surface. Its energy loss due to air absorption can also be included for accurate simulations.[3,23] Whenever a particle passes through a detector volume, whose center position can be considered as the location of a point detector, the energy carried by this particle is recorded versus its arrival time. The recorded energy is proportional to the distance traveled by the ray particle within the detector.[3] By adding up energies contributed by the particles passing through the detector volume with a specified time resolution, the energy impulse response function can be computed. The computer algorithms which were used in this work are based on those presented by Kulowski.[2]

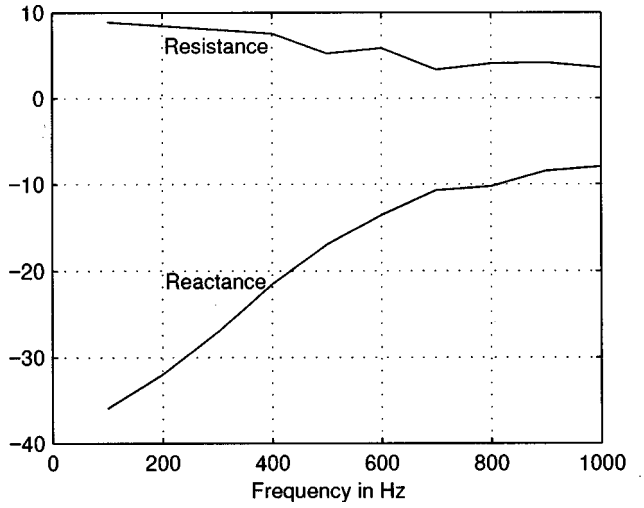


FIG. 15. The normal specific acoustic impedance ξ of the carpet.

TABLE II. The source and microphone positions (in meters) for measurement 1.

Source position	Microphone position
(5.30, 2.10, 0.78)	(1.50, 3.10, 1.65)

B. Image model

The monopole source and all the contributing image sources radiate spherically a certain amount of energy or power at one time instant $t=0$. Usually only a small portion of the vast number of all the possible image sources contribute at a particular detector position. Sifting out only the contributing image sources can give rise to an enormously long computation time. This can be overcome by using the coordinate transformation method[6] or the ray tracing technique.[7] In the work described here, an algorithm specified by Vorländer[7] was used. After reflection from a surface the sound energy from a source is decreased by the diffuse absorption coefficient of the surface. Due to the spherical radiation the detected energy at a point detector is inversely proportional to the squared distance between the source and detector. For example, the contribution to the energy impulse response made by an *n*th order image source, when the distance between this image source and the detector is r_n, can be expressed by

$$\frac{(1-\alpha_1)(1-\alpha_2)\cdots(1-\alpha_n)}{4\pi r_n^2}, \tag{14}$$

where the energy arrival time is r_n/c. By adding up energies at a detector position with a specified time resolution, the energy impulse response function can be constructed.

C. Phase image model

The physical principles are the same as those explained in the previous section for the single reflecting surface problem, and the algorithmic principles are the same as those for the conventional image model. For example, the contribution to the total complex acoustic pressure made by an *n*th order image source can be expressed by

$$p_n(\omega)=\frac{\rho_0 R_1(\theta_1,\omega)R_2(\theta_2,\omega)\cdots R_n(\theta_n,\omega)}{4\pi r_n}\dot{q}(\omega)e^{-jkr_n}. \tag{15}$$

The total acoustic pressure is then computed from

$$p(\omega)=\frac{\rho_0\dot{q}(\omega)e^{-jkr_0}}{4\pi r_0}+\sum_{n=1}^{N} p_n(\omega). \tag{16}$$

TABLE III. The source and microphone positions (in meters) for measurement 2.

Source position	Microphone position
(5.30, 1.10, 0.78)	(1.50, 2.10, 1.65)

TABLE IV. The program execution time on the computer for three models.

Phase image model	Ray tracing model	Image model
40 mins	19 s	25 s

The transient response can be obtained by using the same technique described in the previous section. The energy impulse response can be readily computed by squaring the pressure impulse response, and then by integration with a specified time resolution.

One penalty for using this model is the longer computation time due to the manipulation of complex numbers instead of real numbers since phase information is used to deal with wave interference. For instance, for a 10th-order image source, 10 complex reflection coefficients must be multiplied together. This also has to be repeated at other frequencies as well. These computational procedures have to be carried out for every individual image source.

III. COMPARISON OF ROOM IMPULSE RESPONSES BETWEEN MEASUREMENTS AND GEOMETRICAL ROOM ACOUSTIC MODELS

Some measurements were made in two rooms: one is a medium-sized room and the other a small room. The pressure responses obtained from measurements are compared with those from the phase image model. The energy impulse responses obtained from measurements are compared with those from the phase image model and the conventional model. The comparisons are made during the initial 100 ms.

The accuracies of the energy impulse responses simulated from models can be assessed by the error measure

$$\left(\sum_{n=0}^{50} |e_{\text{meas}} - e_{\text{model}}| \Big/ \sum_{n=0}^{50} e_{\text{meas}} \right) \times 100(\%), \qquad (17)$$

where $e_{\text{meas}}(n)$ and $e_{\text{model}}(n)$ are the measured and modeled energy impulse responses at the nth time interval of 1-ms duration, and the initial 50 ms were considered.

A. Medium-sized room

The configuration of the room is drawn in Fig. 14(a) and (b). The monopole source used is the same as that shown in Fig. 4. The source strength signal chosen is the same as that which is plotted in Fig. 12. The floor is covered with carpet. One vertical surface ($x=0$) is concrete. The other wall surfaces are plaster. The average reverberation time of this room is about 800 ms in the 500-Hz octave band. The normal specific acoustic impedance ξ of the carpet was measured and plotted in Fig. 15. The diffuse absorption coefficients of the concrete and plaster walls were assumed to be 0.02 and 0.07, respectively, and to be independent of frequency. The real parts of their normal reflection coefficients were assumed to be $0.99(=\sqrt{1-0.02})$ and $0.96(=\sqrt{1-0.07})$, respectively, and to be independent of frequency. Also their imaginary parts were assumed to be zero since the concrete and plaster walls are acoustically hard. The results of two

TABLE V. The error measures of the simulated energy impulse responses in percentage (%).

	Measurement 1	Measurement 2
Phase image model	45	35
Conventional model	165	175

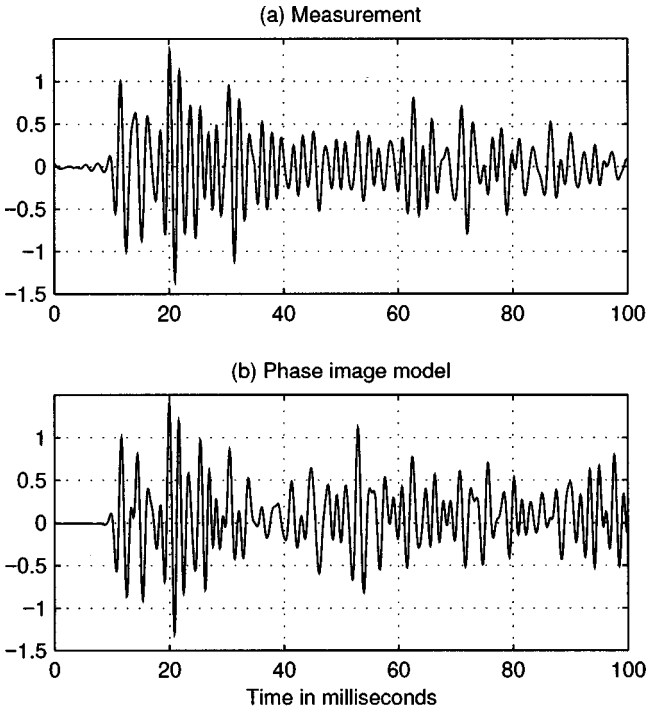


FIG. 16. Comparison of Gabor pulse responses. (Measurement 1 of medium-sized room.)

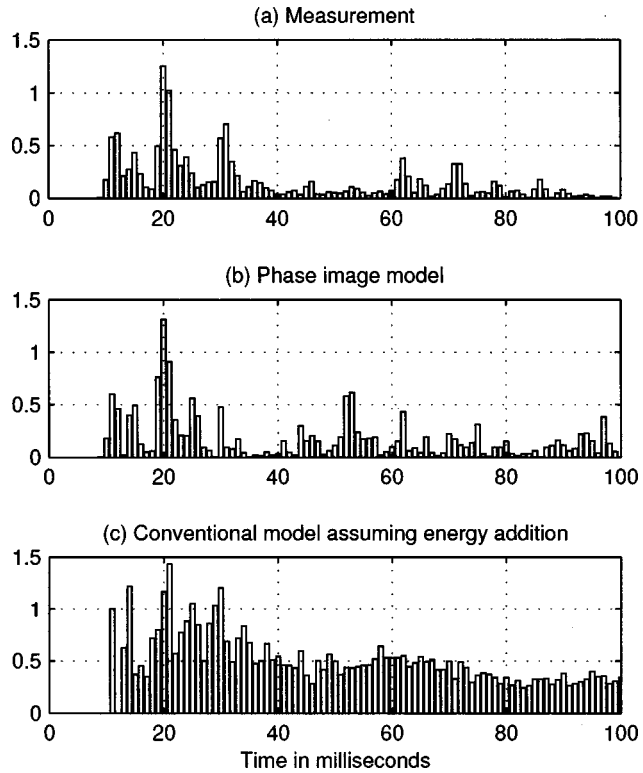


FIG. 17. Comparison of energy impulse responses. (Measurement 1 of medium-sized room.)

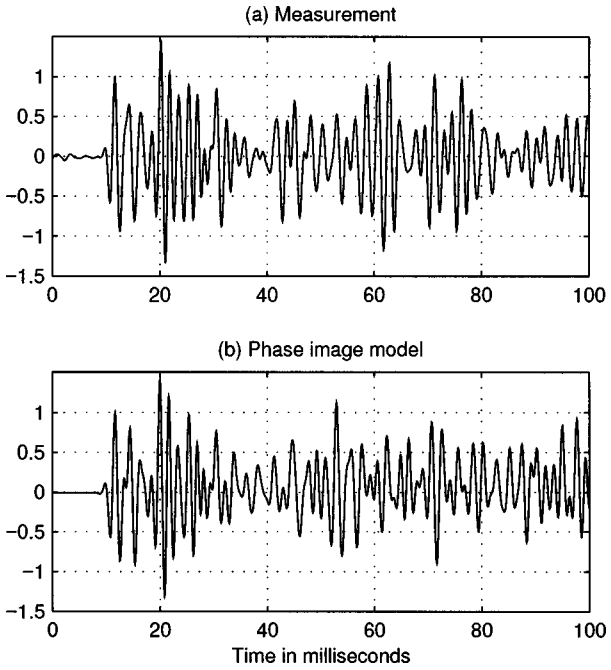


FIG. 18. Comparison of Gabor pulse responses. (Measurement 2 of medium-sized room.)

measurements are presented in this paper. The source and microphone positions for each measurement are tabulated in Tables II and III. The execution time on the computer for the three models are listed in Table IV. The accuracies of the simulated energy impulse responses are listed in Table V. The results are shown in Figs. 16–19.

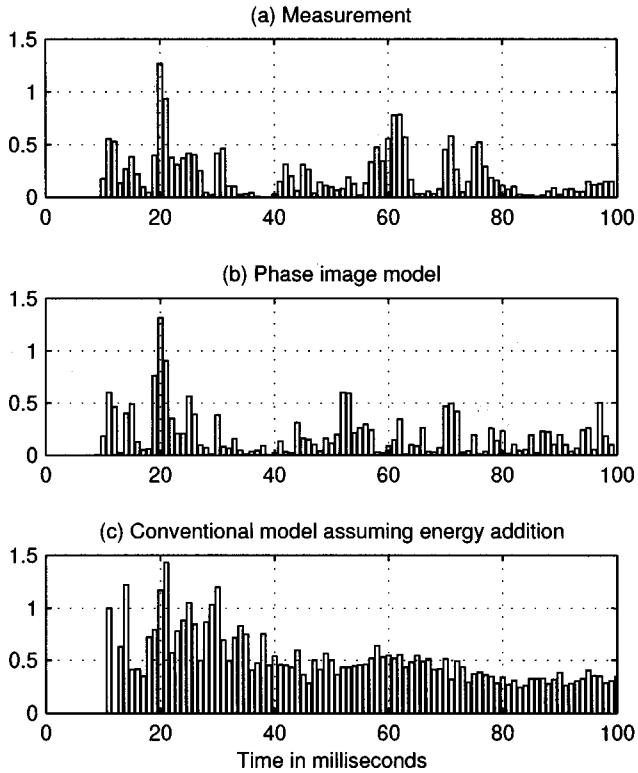


FIG. 19. Comparison of energy impulse responses. (Measurement 2 of medium-sized room.)

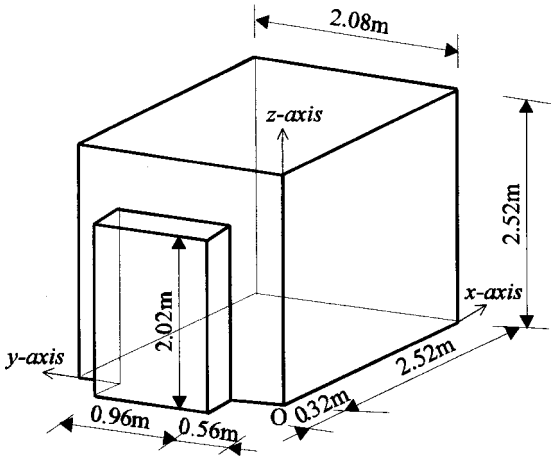


FIG. 20. Configuration of small room. O is the origin of the coordinate system.

B. Small room

The configuration of the room is drawn in Fig. 20. The source strength signal chosen is plotted in Fig. 21. The monopole source used in the measurements is shown in Fig. 22. The results of three measurements are presented in this paper. For measurement 1 and 2, the room is covered with two acoustic materials: The floor is covered with plastic foam and the other wall surfaces are plaster. The average reverberation time is 450 ms in the 2-KHz octave band. The foam is an open cell type and its structure is less dense than that used for the single reflecting surface measurement. Its specific acoustic impedance is plotted in Fig. 23. For measurement 3, one vertical surface ($x=2.52$; refer to Fig. 20) is additionally covered with an insulation board. Its specific acoustic impedance is plotted in Fig. 24. The average rever-

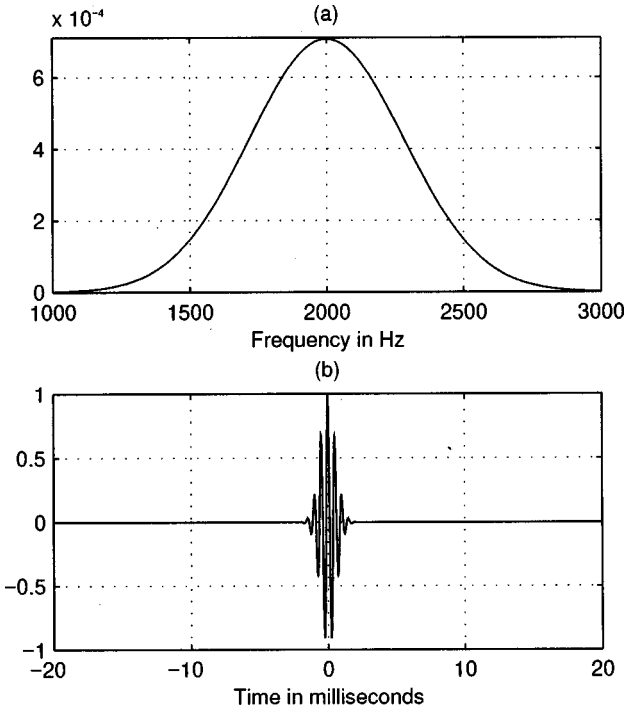


FIG. 21. The source strength signal with $\Delta t \cong 1/1000$ s, $\Delta f \cong 500$ Hz, $f_0 = 2000$ Hz. (a) $\phi_2(f)$, and (b) $\psi_2(t)$. Refer to Eqs. (9) and (12).

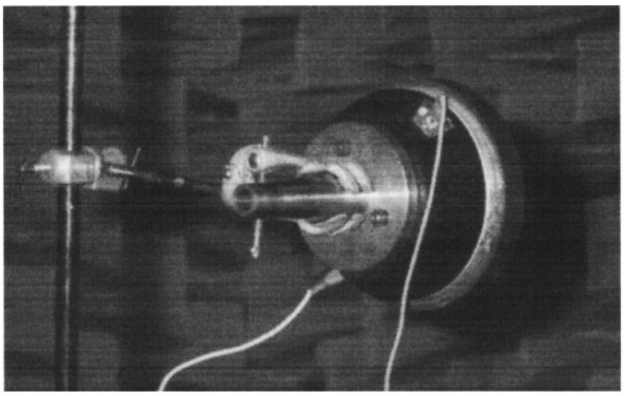


FIG. 22. Monopole source used in small room measurement. The opening diameter of the horn is 13 mm. A tweeter unit is fixed between the horn and the cylindrical body. (Courtesy of Dr. K. R. Holland and Professor F. J. Fahy, University of Southampton.)

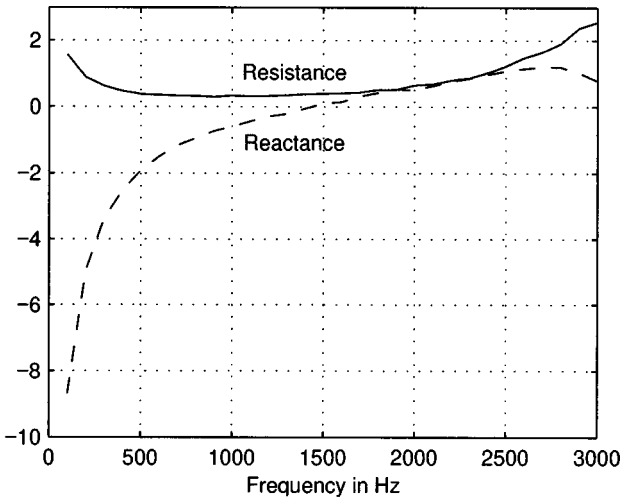


FIG. 23. The normal specific acoustic impedance ξ of the plastic foam.

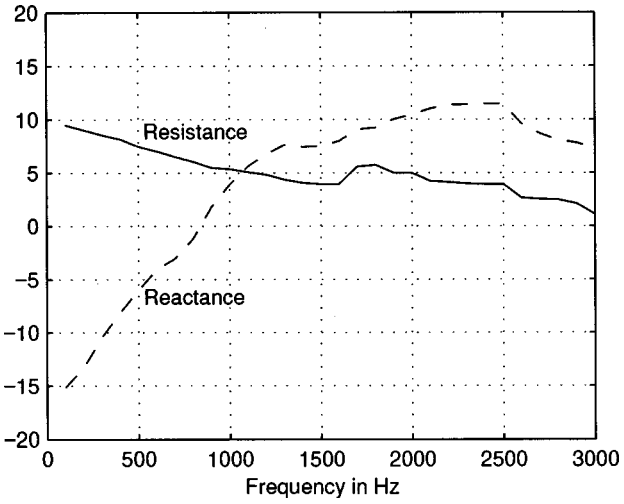


FIG. 24. The normal specific acoustic impedance ξ of the insulation board.

TABLE VI. The source and microphone positions (in meters) for measurements 1 and 3.

Source position	Microphone position
(0.45, 0.34, 0.40)	(1.58, 1.08, 1.55)

TABLE VII. The source and microphone positions (in meters) for measurement 2.

Source position	Microphone position
(0.74, 0.91, 0.95)	(2.09, 1.41, 1.30)

TABLE VIII. The program execution time on the computer for three models.

Phase image model	Ray tracing model	Image model
90 min	1 min 50 s	2 min 16 s

TABLE IX. The error measures of the simulated energy impulse responses in percentage (%).

	Measurement 1	Measurement 2	Measurement 3
Phase image model	71	55	56
Conventional model	77	70	62

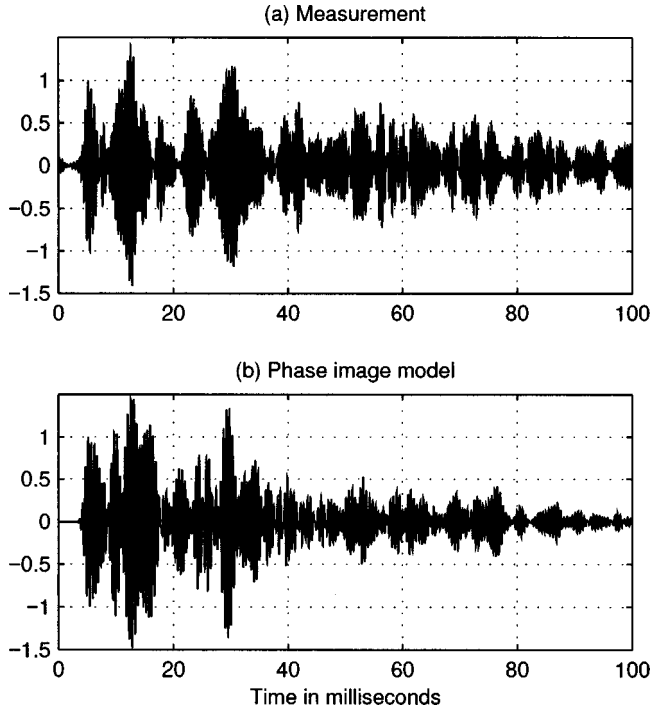


FIG. 25. Comparison of Gabor pulse responses between measurement and phase image model. (Measurement 1 of small room.)

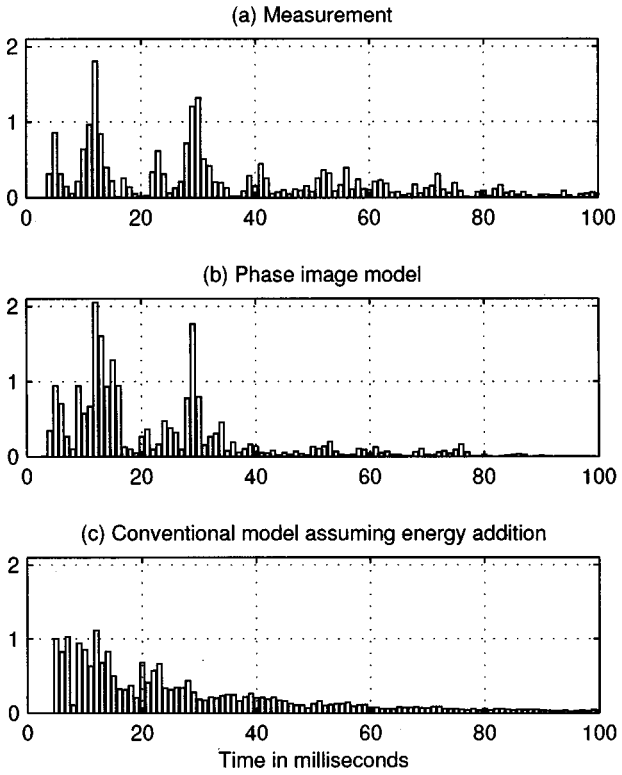


FIG. 26. Comparison of energy impulse responses. (Measurement 1 of small room.)

beration time is 350 ms in the 2-KHz octave band. The diffuse absorption coefficient of the plaster walls was assumed to be 0.04, and to be independent of frequency. The real part of its normal reflection coefficient was assumed to be 0.98 $(=\sqrt{1-0.04})$, and to be independent of frequency. Its imaginary part was again assumed to be zero since the plaster walls are acoustically hard. The source and microphone positions for each measurement are tabulated in Tables VI and VII. The execution time on the computer for the three models are listed in Table VIII. The accuracies of the simulated energy impulse responses are listed in Table IX. The results are shown in Figs. 25–30.

IV. THE EFFECT OF PHASE SHIFTS ON REFLECTION

In this section the effect of the phase shift produced by the complex surface impedance is investigated by replacing the surface impedance by a purely positive real value that has the same absorption coefficient as the true surface impedance. The aim of this section is to compare the predictions from the phase image model for both cases: when complex impedance values are used and when positive real impedance values are used. Again the reflecting surfaces are assumed to be locally reacting. The positive real impedance which has the same absorption coefficient as the true complex impedance can be evaluated from the normal absorption coefficient α_n. The positive real normal reflection coefficient R_n can be approximated by

$$R_n=\sqrt{1-\alpha_n}. \tag{18}$$

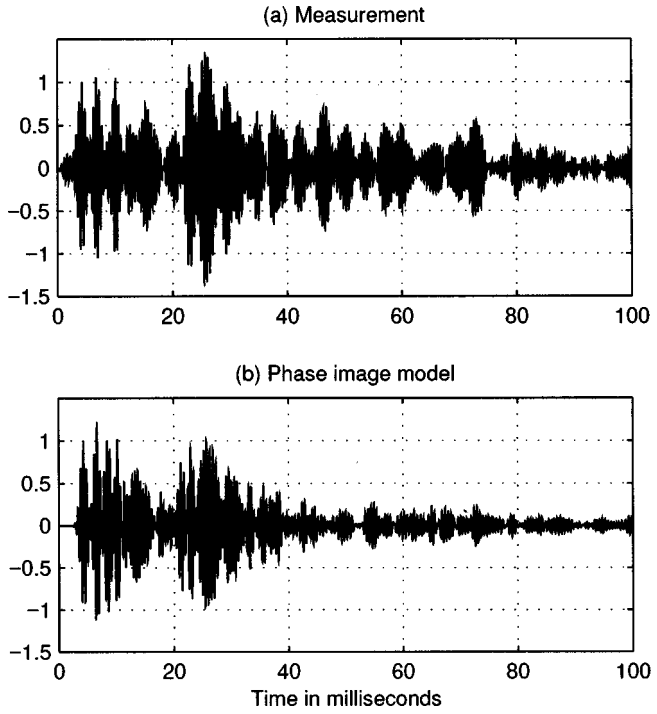


FIG. 27. Comparison of Gabor pulse responses between measurement and phase image model. (Measurement 2 of small room.)

When R_n is substituted into Eq. (3), the positive real impedance, ξ_{pos} will be given by

$$\xi_{\text{pos}}=\frac{1+R_n}{1-R_n}. \tag{19}$$

The configuration of the room is the same as that shown in Fig 14(b). The source and receiver positions are defined at (5.30 m, 2.10 m, 0.78 m) and (1.50 m, 2.10 m, 1.65 m), respectively. The whole surface is covered with a single acoustic material. Its specific acoustic impedance is the same as that which is plotted in Fig. 8. The frequency response functions are compared in Fig. 31. The corresponding response to the Gabor pulse shown in Fig. 12 is also shown in Fig. 32. There is a clear difference between the frequency response functions and Gabor pulse responses in this case which shows that, at least for fairly absorbent surfaces, the complex impedance may be a significant factor in determining the accuracy of the model. It is also time to say that with less absorbing surfaces, the effect of the complex reflections was found to be far less significant.

It is worthwhile to mention previous work which suggests that any two sound signals having the same amplitude spectra will convey almost identical aural impressions regardless of the differences between their phase spectra.[15] If the two frequency response functions compared in Fig. 31 have similar amplitude spectra, they may convey similar aural impressions. However, their differences in aural impressions should really be evaluated by subjective listening tests before any definitive conclusions can be reached. The task of subjective listening tests is beyond the scope of this work and is clearly in need of future investigation.

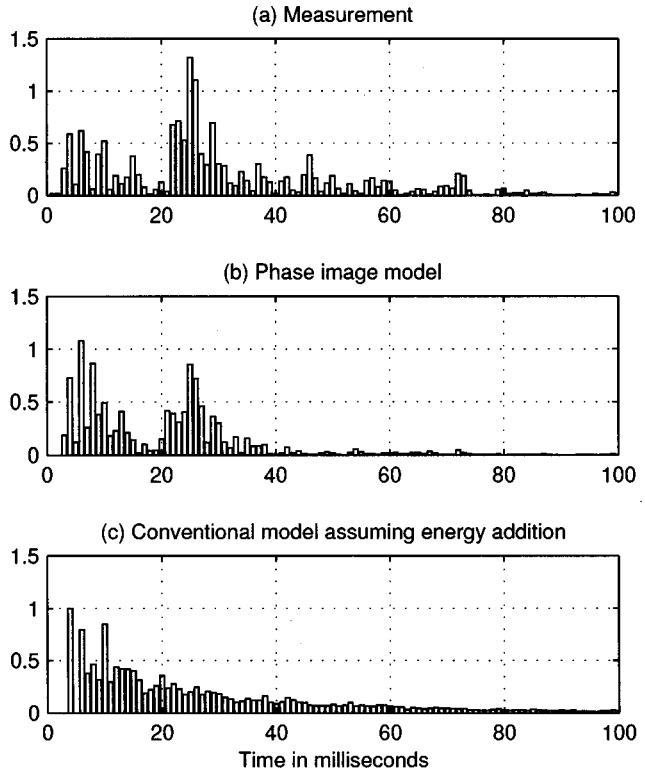


FIG. 28. Comparison of energy impulse responses. (Measurement 2 of small room.)

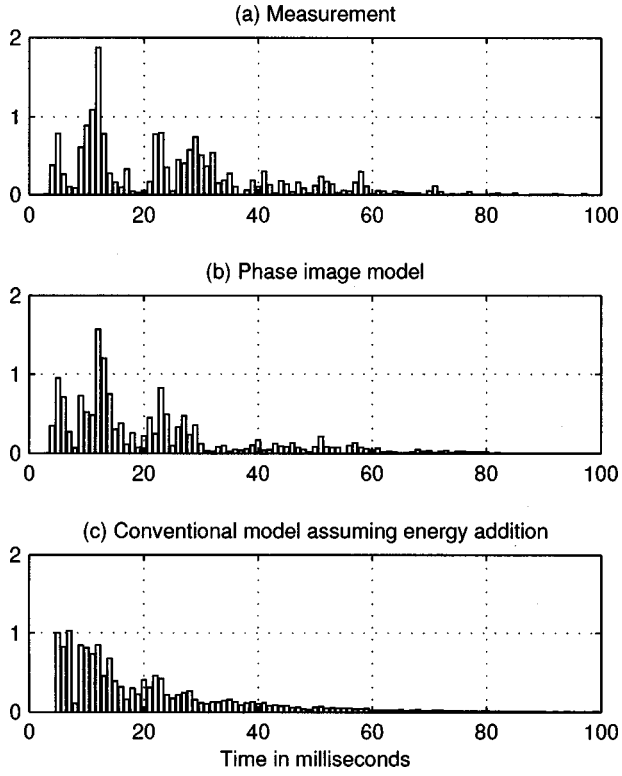


FIG. 30. Comparison of energy impulse responses. (Measurement 3 of small room.)

V. CONCLUSION

It has been demonstrated by measurements that the important physical phenomenon of interference occurring in room acoustic fields can be reasonably well modeled by the phase image model. The conventional model, however,

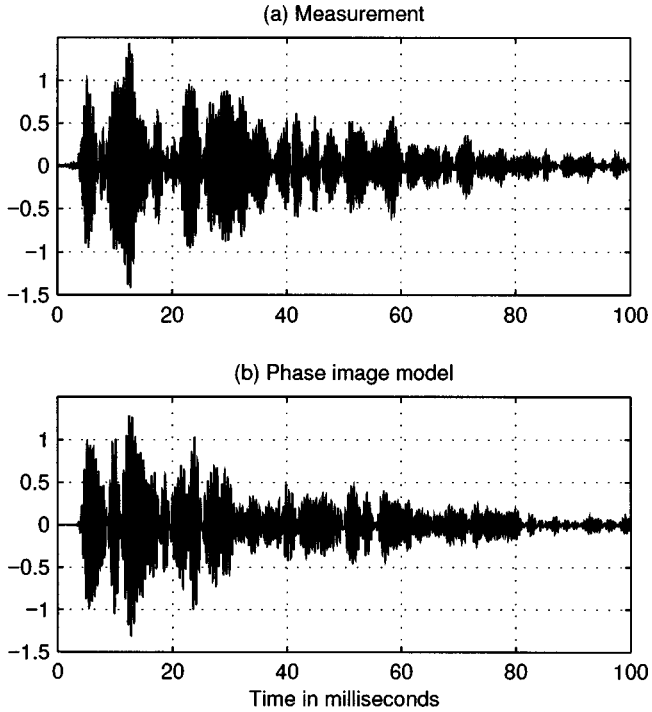


FIG. 29. Comparison of Gabor pulse responses between measurement and phase image model. (Measurement 3 of small room.)

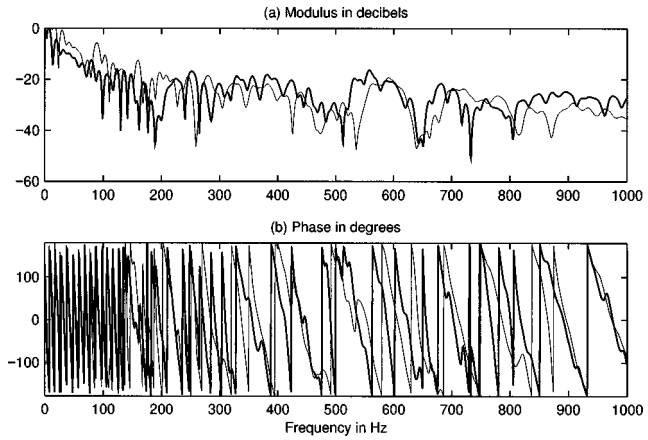


FIG. 31. Comparison of frequency response functions. Complex impedance: Thick solid line; Positive real impedance: Thin solid line.

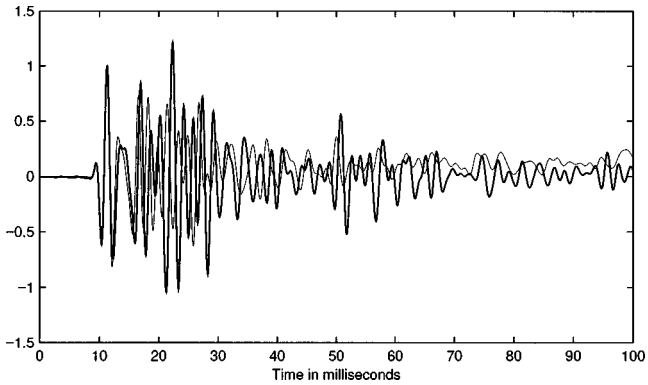


FIG. 32. Comparison of Gabor pulse responses. Complex impedance: Thick solid line; Positive real impedance: Thin solid line.

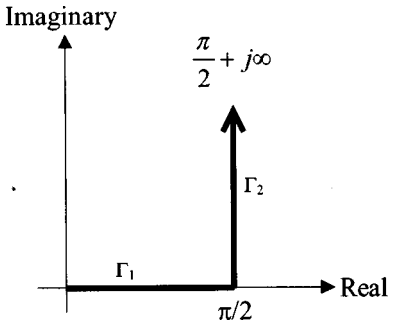


FIG. A1. Integration contour, $\Gamma_\theta(=\Gamma_1+\Gamma_2)$, in the complex θ-plane.

which does not attempt to model interference gives a much poorer representation of the response of rooms to a short duration pulse. By using the Gabor pulse as a source strength signal, it was possible to best resolve the direct and reflected waves in the pressure pulse responses measured in a given frequency band. Five measurements were carried in total for the medium-sized room. The average error measures of the energy impulse responses for the phase and conventional models were about 46% and 173%, respectively. Eight measurements were carried in total for the small room. The average error measures of the energy impulse responses for the phase and conventional models were about 63% and 70%, respectively.

The early part of the pulse response of rooms will have a big impact on aural impressions and will be dominated by specular reflections. Without proper modeling the simulated pulse responses are likely to sound unnatural to human ears. The later part, although diminished in level, will also play an important role in creating aural impressions. This part will be dominated by diffraction and scattering. Currently considerable errort is being made to model these two phenomena. For accurate modeling of room impulse responses, it is very important to develop a rigorous technique for measuring the incident angle dependent reflection coefficients or specific acoustic impedances of acoustic materials.[25]

ACKNOWLEDGMENTS

The authors gratefully acknowledge Dr. K. R. Holland and Professor F. J. Fahy for their helpful advice on measurement techniques.

APPENDIX: BASIC THEORY FOR NUMERICAL INTEGRATION OF EQ. (2)

The integration contour Γ_θ can be divided into two paths. Path Γ_1 extends from 0 to $\pi/2$ on the real axis. Path Γ_2 extends from $\pi/2$ to $\pi/2+j\infty$. This is illustrated Fig. A1. Equation (2) is also known as the Sommerfeld Integral.

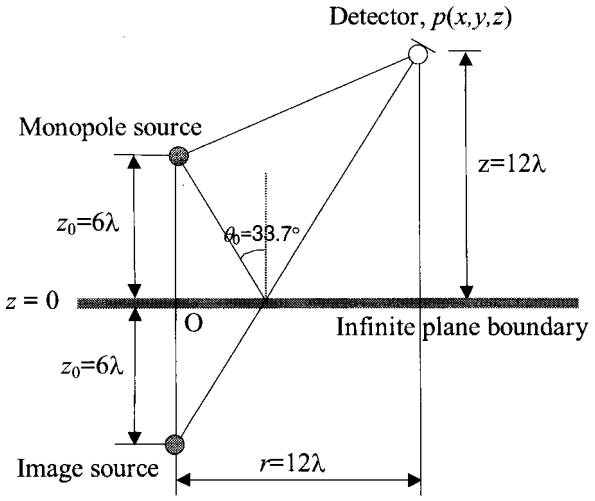


Fig. A2. Configuration of source and detector in a half-space. O is the origin of the coordinate system.

The Sommerfeld Integral evaluated along path Γ_1, which will be denoted by $p_{\text{refl},1}$, will be given by

$$p_{\text{refl},1}=-\frac{j\rho_0 k}{4\pi}\int_0^{\pi/2} J_0(kr\sin\alpha)\exp(-jk(z+z_0)\cos\alpha)\times R(\alpha)\sin\alpha\, d\alpha, \tag{A1}$$

where $R(\alpha)$ is

$$R(\alpha)=\frac{\xi\cos\alpha-1}{\xi\cos\alpha+1}. \tag{A2}$$

By using the identity

$$\cos\left(\frac{\pi}{2}+j\beta\right)=\cos\frac{\pi}{2}\cosh\beta-j\sin\frac{\pi}{2}\sinh\beta=-j\sinh\beta,$$
$$\sin\left(\frac{\pi}{2}+j\beta\right)=\sin\frac{\pi}{2}\cosh\beta+j\cos\frac{\pi}{2}\sinh\beta=\cosh\beta. \tag{A3}$$

The Sommerfeld integral evaluated along path Γ_2, which will be denoted by $p_{\text{refl},2}$, will be given by

$$\sin\left(\frac{\pi}{2}+j\beta\right)=\sin\frac{\pi}{2}\cosh\beta+j\cos\frac{\pi}{2}\sinh\beta=\cosh\beta, \tag{A4}$$

where $R((\pi/2)+j\beta)$ is

$$R\left(\frac{\pi}{2}+j\beta\right)=\frac{j\xi\sinh\beta+1}{j\xi\sinh\beta-1}. \tag{A5}$$

Thus the solution of the Sommerfeld integral is the sum of $p_{\text{refl},1}$ and $p_{\text{refl},2}\cdot p_{\text{refl},1}$ is a definite integral but $p_{\text{refl},2}$ is an improper integral whose integration variable β extends from 0 to ∞. Except grazing incidence, it can be easily noticed that

TABLE AI. Comparison of the numerical integrations with different spacings of $\Delta\alpha$ and $\Delta\beta$.

	$p_{\text{refl},1}$	$p_{\text{refl},2}$	$p_{\text{refl}}=p_{\text{refl},1}+p_{\text{refl},2}$
$\Delta\alpha=\Delta\beta=10^{-4}$	$-1.003\,23e-2+j2.06514e-2$	$-2.48551e-3-j3.002\,00e-4$	$-1.25178e-2+j2.035\,12e-2$
$\Delta\alpha=\Delta\beta=10^{-5}$	$-1.003\,23e-2+j2.06514e-2$	$-2.48550e-3-j3.002\,02e-4$	$-1.25178e-2+j2.035\,12e-2$
$\Delta\alpha=\Delta\beta=10^{-6}$	$-1.003\,23e-2+j2.06514e-2$	$-2.48550e-3-j3.002\,02e-4$	$-1.25178e-2+j2.035\,12e-2$

the integrand of Eq. (A4) is a rapidly decaying function of β converging toward zero. Let us express the integrand of Eq. (A4) by $F(\beta)$. Then $p_{\text{refl},2}$ can be expressed by

$$p_{\text{refl},2}=\frac{\rho_0 k}{4\pi}\int_0^1 F(\beta)d\beta+\frac{\rho_0 k}{4\pi}\int_1^2 F(\beta)d\beta + \frac{\rho_0 k}{4\pi}\int_2^3 F(\beta)d\beta+\cdots. \tag{A6}$$

Let us denote each term in Eq. (A6) by $A(0)$, $A(1)$, $A(2)$, $A(4),\cdots$. Then Eq. (A6) can be represented by

$$p_{\text{refl},2}=A(0)+A(1)+A(2)+\cdots. \tag{A7}$$

If the series $A(0),A(1),A(2),\cdots$ is rapidly converging, then $p_{\text{refl},2}$ can be computed. Actually they do converge except grazing incidence. The numerical technique used is the Trapezoidal rule. $d\alpha$ and $d\beta$ in Eqs. (A1) and (A4) will be replaced by $\Delta\alpha$ and $\Delta\beta$, which are constant steps on α and β axis, respectively. For good accuracy, $\Delta\alpha$ and $\Delta\beta$ should be made as small as possible. One example of the numerical integration was undertaken for the following configuration shown in Fig. A2. Several values of $\Delta\alpha$ and $\Delta\beta$ were used in order to investigate the convergence of Eqs. (A1) and (A4). The specific acoustic impedance ξ of the surface was assumed to be $5.00-j11.00$ with local reaction assumed. The values of c and ρ_0 used are 344 m/s and 1.21 kg/m^3.

The results are presented in Table AI. It is evident that as $\Delta\alpha$ and $\Delta\beta$ become smaller, the Sommerfeld integral converges.

[1] A. Krokstad *et al.*, ''Calculating the acoustical room response by the use of a ray tracing technique,'' J. Sound Vib. **8**, 118–125 (1968).
[2] A. Kulowski, ''Algorithmic representation of the ray tracing technique,'' Appl. Acoust. **18**, 449–469 (1985).
[3] A. M. Ondet and J. L. Barby, ''Modeling of sound propagation in fitted workshop using ray tracing,'' J. Acoust. Soc. Am. **85**, 787–796 (1989).
[4] J. B. Allen and D. A. Berkeley, ''Image method for efficiently simulating small-room,'' J. Acoust. Soc. Am. **65**, 943–950 (1978).
[5] J. Borish, ''Extension of the image model to arbitrary polyhedra,'' J. Acoust. Soc. Am. **75**, 1827–1836 (1984).
[6] H. Lee and B. Lee, ''An efficient algorithm for the Image Model Technique,'' Appl. Acoust. **24**, 87–115 (1988).
[7] M. Vorländer, ''Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm,'' J. Acoust. Soc. Am. **86**, 172–178 (1989).
[8] R. Heinz, ''Binaural spatial simulation using a combined method—Separate simulation of the geometrical and the diffuse components,'' Acustica **79**, 207–220 (1993).
[9] B. Dalenbäck, ''Room acoustic prediction and auralization based on a unified treatment of diffuse and specular reflection,'' 15th ICA, Trondheim, Norway, pp. 425–428 (1995).
[10] M. J. Hodgson, ''Evidence of diffuse surface reflections in rooms,'' J. Acoust. Soc. Am. **89**, 765–771 (1991).
[11] E. Mommertz and M. Vorländer, ''Measurement of scattering coefficients of surfaces in the reverberation chamber and in the free field,'' 15th ICA, Trondheim, Norway, pp. 577–580 (1995).
[12] U. E. Stephenson, ''Quantized pyramidal beam tracing—A new algorithm for room acoustics and noise immission prognosis,'' Acustica **82**, 517–525 (1996).
[13] M. Tran Van Nhieu, ''Diffraction by the edge of a three-dimensional object,'' J. Acoust. Soc. Am. **99**, 79–87 (1996).
[14] F. Mortessagne and O. Legrand, ''Semi-classical acoustic time response in 2-D rooms,'' Acustica **82**, 152 (1996).
[15] H. Kuttruff, ''On the audibility of phase distortions in rooms and its significance for sound reproduction and digital simulation in room acoustics,'' Acustica **74**, 3–7 (1991).
[16] Y. W. Lam, ''A comparison of three diffuse reflection modeling methods used in room acoustics computer models,'' J. Acoust. Soc. Am. **100**, 2181–2192 (1996).
[17] E. De Geest and H. Pätzold, ''Comparison between room transmission functions calculated with a boundary element method and a ray tracing method including phase,'' Proceedings of Internoise **96**, 3177–3180 (1996).
[18] D. Gabor, ''Theory of Communication,'' Journal of the IEEE, London **93(III)**, 429–457 (1946).
[19] J. K. Hammond and P. R. White, ''The analysis of nonstationary signals using time-frequency methods,'' J. Sound Vib. **190**, 419–447 (1996).
[20] L. M. Brekhovskikh, *Waves in Layered Media*, 2nd ed. (Academic, New York, 1980).
[21] S. Thomasson, ''Reflection of waves from a point source by an impedance boundary,'' J. Acoust. Soc. Am. **59**, 780–785 (1976).
[22] G. Wills, ''Shallow water propagation from an embedded source,'' Ph.D. thesis, The Institute of Sound and Vibration Research, University of Southampton (1996).
[23] H. Kuttruff, *Room Acoustics*, 3rd ed. (Elsevier, Applied Science, 1991).
[24] F. S. Crawford, *Waves*, Berkeley Physics Course—Vol. 3, Education Development Center), 3rd ed. (1968), pp. 286–292.
[25] B. Brouard, D. Lafarge, J. F. Allard, and M. Tamura, ''Measurement and prediction of the reflection coefficient of porous layers at oblique incidence and for inhomogeneous waves,'' J. Acoust. Soc. Am. **99**, 100–107 (1996).

Radius of curvature estimation and localization of targets using multiple sonar sensors

Billur Barshan
Department of Electrical Engineering, Bilkent University, Bilkent, 06533 Ankara, Turkey

Ali Şafak Sekmen
Center for Intelligent Systems, Department of Electrical and Computer Engineering, Vanderbilt University, Box 1824, Station B, Nashville, Tennessee 37235

(Received 12 August 1998; accepted 31 December 1998)

Acoustic sensors have been widely used in time-of-flight ranging systems since they are inexpensive and convenient to use. One of the most important limitations of these sensors is their low angular resolution. To improve the angular resolution and the accuracy, a novel, flexible, and adaptive three-dimensional (3-D) multi-sensor sonar system is described for estimating the radius of curvature and location of cylindrical and spherical targets. Point, line, and planar targets are included as limiting cases which are important for the characterization of typical environments. Sensitivity analysis of the curvature estimate with respect to measurement errors and certain system parameters is provided. The analysis and the simulations are verified by experiments in 2-D with specularly reflecting cylindrical and planar targets, using a real sonar system. Typical accuracies in range and azimuth are 0.18 mm and 0.1°, respectively. Accuracy of the curvature estimation depends on the target type and system parameters such as transducer separation and operating range. The adaptive configuration brings an improvement varying between 35% and 45% in the accuracy of the curvature estimate. The presented results are useful for target differentiation and tracking applications. © *1999 Acoustical Society of America.* [S0001-4966(99)03904-1]

PACS numbers: 43.58.−e, 43.28.Tc, 43.60.Qv, 43.35.Yb [SLE]

INTRODUCTION

Ultrasonic transducers are a convenient and inexpensive means for intelligent systems to build models of their environment. However, these sensors are limited by their wide beamwidth which makes accurate localization of targets difficult. To increase the localization accuracy, an adaptive sensor configuration composed of multiple ultrasonic transducers is proposed that is capable of estimating the radius of curvature and location of spheres, cylinders, point, line, and planar targets. Consequently, these basic types of reflectors can be differentiated.

Target localization has been extensively studied in earlier work. In Ref. 1, time-delay estimation for active/passive localization in underwater sonar is reviewed with references to benchmark work. In particular, ocean effects which require sonar adaptation are considered. Adaptive sonar arrays have been also used by other researchers to add flexibility to their systems.[2] Coherent and incoherent processing techniques of time-delay estimation have been addressed in Refs. 3, 4. Active, wide-band detection and localization of targets in a dense and uncertain multipath environment has been considered in Ref. 5. The review article in Ref. 6 considers numerical schemes for accurate processing of information from both active and passive acoustic arrays.

Sonar sensing has many applications for intelligent systems operating in three-dimensional (3-D) environments, such as airborne or underwater robots. Several researchers have investigated the limitations of sonar for 3-D target recognition, discrimination, and tracking: Self-contained navigation systems have been devised for underwater vehicles, capable of tracking and producing continuous range information from a passive target.[7] In Ref. 8, an approach is described to the construction of 3-D stochastic models for intelligent systems exploring the underwater environment. In Ref. 9, the minimum amount of information and actuation needed to track a ball in 3-D has been determined and implemented using qualitative methods. Hong and Kleeman have investigated the geometry of 3-D corner cubes using a low-sample rate equilateral triangular sonar configuration.[10] Kleeman and Akbarally have classified and discriminated the target primitives commonly occurring in 3-D space.[11] Peremans *et al.*[12] and Sabatini[13,14] both have investigated curved reflectors using linear array configurations. In Ref. 15, an analytical approach to surface curvature extraction is described which employs ultrasonic echo trajectories and differential geometry. In Refs. 16 and 17, binaural sonar information is fused for accurate object recognition using a system which adaptively changes its position and configuration in response to the echoes it detects. Curvature estimation has been also important in image analysis to provide viewpoint-independent cues for shape classification.[18]

Some sonar systems attempt to emulate the remarkable perception and pattern recognition capabilities of bats and dolphins in extracting detailed information about their environments from acoustic echo returns.[19–21] Artificial neural networks have been widely used for this purpose, to process time and/or frequency representations of sonar echo signals. For example, one application is in the classification of sonar returns from undersea targets where the targets may be made of different materials, have different shape, buried in mud or sediment, or exist in the presence of other reflectors in the

environment.[21–23] In Ref. 24, cylinder-wall thicknesses discriminating capability of artificial neural networks is compared to that of dolphins. In Ref. 25, artificial neural networks are applied to classifying underwater active sonar returns with different numbers of peaks. Another system can recognize 3-D cubes and tetrahedrons, independent of their orientation with the help of neural networks.[19]

Acoustic imaging of extended targets by means of synthetic-aperture sonar has been considered in Ref. 26, where echoes from spherical and cylindrical targets laid down on a seabed are processed together with random echoes from the sea bottom. In Ref. 27, Stergiopoulos reviews the implementation of adaptive synthetic-aperture processing schemes in integrated active–passive sonar systems.

In this paper, an adaptive sonar configuration is used for radius of curvature estimation and localization of targets. When the reflection point of the target is not along the line-of-sight of the transducer, the amplitude of the reflected signal is smaller, which decreases the signal-to-noise ratio (SNR) and worsens the accuracy. To reduce this effect, the transducers are rotated toward the target to obtain more nearly accurate estimates.

The organization of the paper is as follows: In Sec. I, background information on acoustic reflection and signal models of sonar sensors is reviewed and motivation for the adaptive configuration is provided. Methods for time-of-flight estimation are discussed in Sec. II. In Sec. III, the geometry of reflection from spherical targets is considered and analyzed for radius of curvature estimation. The important limiting cases of point and planar targets are highlighted. Sensitivity analysis of curvature estimation is provided with respect to measurement errors and variations in some of the system parameters in Sec. IV. Section V presents the simulation results. A detailed description of the sensing device used in this study is provided in Sec. VI A. Experimental results which verify the analysis and the simulations are presented in Sec. VI B. Section VII briefly discusses the use of the method for target differentiation. Finally, conclusions are drawn and directions for future work are motivated.

I. ACOUSTIC REFLECTION AND SIGNAL MODELS

The characteristics of the radiation pattern of an acoustic transducer are different in the near-field (or Fresnel) region and the far-field (or Fraunhofer) region.[28] In this study, assuming all targets of interest are located in the far field, the far-field model of a piston-type transducer having a circular aperture is used.[28] For a *single* frequency of excitation, the far-field characteristics at range z and angular deviation α from the line-of-sight are described by[29,30]

$$p_{z,\alpha}=\frac{p_{\max}z_{\min}}{z}\,\frac{J_1(ka\sin\alpha)}{ka\sin\alpha}\quad \text{for } z\geqslant z_{\min}, \tag{1}$$

where $J_1(.)$ is the Bessel function of the first order of the first kind and $p_{\max}$ is the propagation pressure amplitude on the beam axis at range $z_{\min}$ along the line-of-sight. $z_{\min}\cong a^2/\lambda$ is the distance at which the far-zone characteristics begin. Although the 2-D cross-section of the characteristics is given here, in fact, the pattern is rotationally symmetric about the line-of-sight.

The half beamwidth angle α_0 in the far-field corresponds to the first zero of the Bessel function in Eq. (1) which occurs at $ka\sin\alpha_0=1.22\pi$, resulting in:[31]

$$\alpha_0=\sin^{-1}\left[\frac{0.61\lambda}{a}\right], \tag{2}$$

where $\lambda=c/f_0$ is the wavelength (f_0 is the resonance frequency of the transducer) and a is the transducer aperture radius.

Since a range of frequencies around f_0 are transmitted, the corresponding beam patterns are superposed and the resulting pattern can be approximated by a Gaussian function centered at zero with standard deviation $\sigma_\alpha=\alpha_0/2$:[32]

$$\tilde{p}_{z,\alpha}=\frac{p_{\max}z_{\min}}{z}e^{-\alpha^2/2\sigma_\alpha^2}\quad \text{for } z\geqslant z_{\min}. \tag{3}$$

For a rigid cylindrical target of infinite height at range z and making an angle α with the line-of-sight of the transducer, the received time signal can be modeled by:[33]

$$s_{z,\alpha}(t)=\rho_c\frac{A_{\max}z_{\min}^{3/2}}{z^{3/2}}e^{-\alpha^2/\sigma_\alpha^2}\;e^{-[t-(t_0+\Delta t_c)]^2/2\sigma_t^2}\times\sin[2\pi f_0(t-t_0)]\quad \text{for } z\geqslant z_{\min}, \tag{4}$$

where ρ_c is the reflection coefficient which increases with the radius of curvature,[34] $A_{\max}$ is the maximum signal amplitude, z is the distance between the transducer and the object surface, t_0 is the time-of-flight, Δt_c is the time difference between the center of the Gaussian window and t_0, and $\sigma_t=1/f_0$. Basically, the received signal envelope has been modeled as a Gaussian function centered at $t_0+\Delta t_c$ with suitably chosen variance σ_t^2.[33] More generally, the model

$$s_{z,\alpha}(t)=k(z)e^{-\alpha^2/\sigma_\alpha^2}\;e^{-[t-(t_0+\Delta t_c)]^2/2\sigma_t^2}\times\sin[2\pi f_0(t-t_0)]\quad \text{for } z\geqslant z_{\min} \tag{5}$$

is capable of representing observed signals for a wide variety of target types and locations in the far zone.[33] Here, $k(z)$ incorporates $A_{\max}$ and ρ_c, and is inversely proportional to some power of the range z depending on target type.[35] The inclination angle α from the line-of-sight is related to the target azimuth and elevation angles θ and ϕ by the relation $\alpha=\cos^{-1}(\cos\theta\cos\phi)$.

With a single stationary transducer, it is not possible to estimate the angular position of the target (θ,ϕ) with better resolution than the angular resolution of the transducer which is approximately $2\alpha_0$. When a range reading is received, all that is known is that the object lies somewhere on a spherical cap subtending a cone of half angle α_0 and radius z, centered at the transducer. This is illustrated in Fig. 1 in 2-D for the transducer in the middle. To improve the angular resolution, the present system employs multiple identical acoustic transducers with center-to-center separation d (Fig. 1). Each transducer can operate both as transmitter and as receiver and detect echo signals reflected from targets within its *sensitivity region*. All members of the sensor configuration can detect targets located within the *joint sensitivity re-*

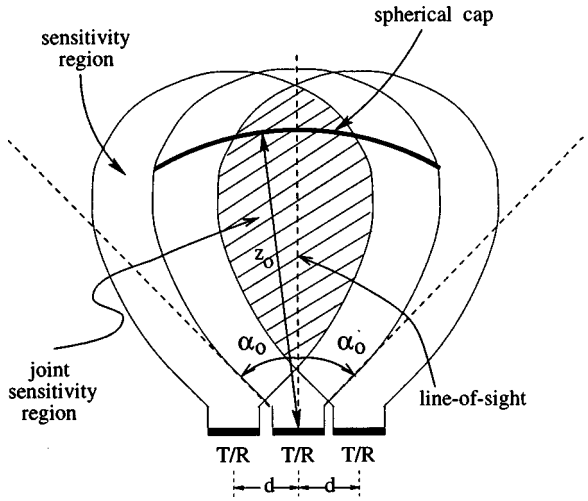


FIG. 1. Minimum distance at which a target is detectable by all three transducers is approximately $[(d-a)/\tan\alpha_0] + (a^2/\lambda)$. This corresponds to the distance between the central transducer and the start of the joint sensitivity region.

gion, which is the overlap of the individual sensitivity regions, as shown in Fig. 1. The extent of this region is different for different targets which, in general, exhibit different reflection properties. For example, for edgelike or polelike targets, this region is much smaller but of similar shape, and for planes, it is more extended.[36]

When the object is not located along the line-of-sight of the transducer (i.e., $\alpha \neq 0°$), there is an exponential decline in the amplitude that decreases the SNR [Eq. (4)]. Hence, information provided by sonar sensors is most reliable when the object lies along the line-of-sight of the transducer. Likewise, at nearby ranges, due to the $k(z)$ term inversely varying with z in Eq. (5), more nearly accurate readings are provided. Therefore, the transducers are rotated around their centers to align the line-of-sight with the object as shown in Fig. 2 in 2-D.

II. TIME-OF-FLIGHT ESTIMATION

The most commonly used sonar ranging systems are based on time-of-flight measurements. In thresholding time-of-flight systems, an echo is produced when a transmitted pulse encounters an object and a range reading $z = ct_0/2$ is produced when the echo amplitude first exceeds a preset

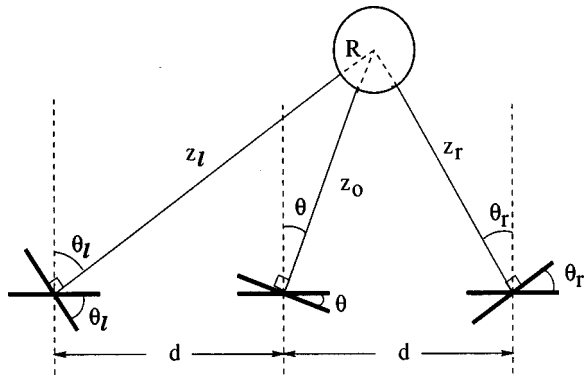


FIG. 2. The spherical object and the initial (flat) and adapted (rotated) sensor configurations illustrated in 2-D.

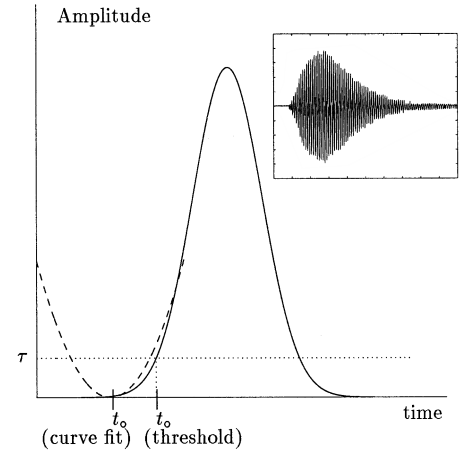


FIG. 3. Envelope of the sonar echo and time-of-flight estimation by thresholding and curve-fitting. Inset: Typical real sonar waveform.

threshold level τ back at the receiver. Here, t_0 is the time-of-flight estimate of the echo signal and c is the speed of sound in air. ($c = 331.4\sqrt{T/273}$ m/s, where T is the absolute temperature in Kelvin. At room temperature, $c = 343.3$ m/s.)

An alternative to simple thresholding is the *curve-fitting* method which is based on fitting a parabola to the onset of the sonar echo in order to reduce the bias on the time-of-flight estimate[35,36] (Fig. 3). In this method, the estimate of t_0, initially obtained by simple thresholding, is further improved by using Levenberg–Marguardt nonlinear least-squares optimization.[37]

Recently, a consolidated comparison of simple thresholding, curve-fitting several variations of the sliding-window used in radar, and cross correlation methods has been provided in terms of their bias, variance, and processing time.[38] The first three methods are much simpler and faster, while offering a variety of attractive compromises between accuracy and system complexity. For instance, curve-fitting improves the bias of the time-of-flight estimate but the variance is comparable and the implementation requires more effort compared to thresholding.

When operated in the pulse-echo mode,[39] the transducers offer complete flexibility in the firing sequence and timing at which the individual transducers can be fired. In the firing pattern used here, each transducer registers only the echo of the signal transmitted by itself. Assuming the target is stationary, the firing is done sequentially to avoid crosstalk between the transducers. This is geometrically simpler to analyze as compared to firing patterns where the transducers simultaneously detect signals transmitted by each other. After each transmission, the detected waveform is recorded and thresholded to obtain a round-trip time-of-flight estimate.

III. TARGET REFLECTION GEOMETRY

In the following analysis, a stationary spherical target of radius R is assumed to be present at spherical coordinates (r, θ, ϕ).

A. Spherical target with radius R

According to the firing pattern described in the previous section, the geometry of Fig. 2 indicates that the noisy distance measurements at the surrounding transducers are:

$$\hat{z}_0=\frac{c\hat{t}_0}{2}=r-R+n_0,$$

$$\hat{z}_r=\frac{c\hat{t}_r}{2}=\sqrt{r^2+d^2-2\,dr\cos\phi\sin\theta}-R+n_r,$$

$$\hat{z}_l=\frac{c\hat{t}_l}{2}=\sqrt{r^2+d^2+2\,dr\cos\phi\sin\theta}-R+n_l, \tag{6}$$

$$\hat{z}_u=\frac{c\hat{t}_u}{2}=\sqrt{r^2+d^2-2\,dr\sin\phi}-R+n_u,$$

$$\hat{z}_d=\frac{c\hat{t}_d}{2}=\sqrt{r^2+d^2+2\,dr\sin\phi}-R+n_d,$$

where $\hat{t}_0$, $\hat{t}_r$, $\hat{t}_l$, $\hat{t}_u$, $\hat{t}_d$ are the time-of-flight measurements at the middle, right, left, up, and down transducers, respectively, n_0, n_r, n_l, n_u, n_d are spatially and temporally uncorrelated zero-mean Gaussian noise. Each measurement confines the possible target locations to a spherical cap defined by the intersection of the sensitivity region and a sphere centered at the corresponding transducer (Fig. 1). At least three measurements are necessary to identify the curvature of the target both in 2-D and 3-D. This is illustrated in Figs. 4 and 5.

In Ref. 12, it has been shown that for the Polaroid transducer, the noise correlation coefficient is negligible since most of the noise on the range measurements is dominated by the thermal noise in the electronics. This is the reason why n_0, n_r, n_l, n_u, and n_d can be modeled as uncorrelated Gaussian noise. Therefore, the error correlation matrix $\mathbf{C}$ and its inverse are approximated as diagonal:

$$\mathbf{C}=\begin{bmatrix}\sigma_{n_0}^2 & 0 & 0 & 0 & 0\\ 0 & \sigma_{n_r}^2 & 0 & 0 & 0\\ 0 & 0 & \sigma_{n_l}^2 & 0 & 0\\ 0 & 0 & 0 & \sigma_{n_u}^2 & 0\\ 0 & 0 & 0 & 0 & \sigma_{n_d}^2\end{bmatrix}, \tag{7}$$

and the conditional probability density function of the measurement vector $\hat{\mathbf{m}}$ is:

$$p(\hat{\mathbf{m}}|r,\theta,\phi,R)=\frac{1}{2\pi|\mathbf{C}|}\exp\left\{-\frac{1}{2}[\hat{\mathbf{m}}-\mathbf{z}(r,\theta,\phi,R)]^{\mathrm{T}}\mathbf{C}^{-1}\times[\hat{\mathbf{m}}-\mathbf{z}(r,\theta,\phi,R)]\right\}, \tag{8}$$

where the vectors $\hat{\mathbf{m}}$ and $\mathbf{z}(r,\theta,R)$ are defined as:

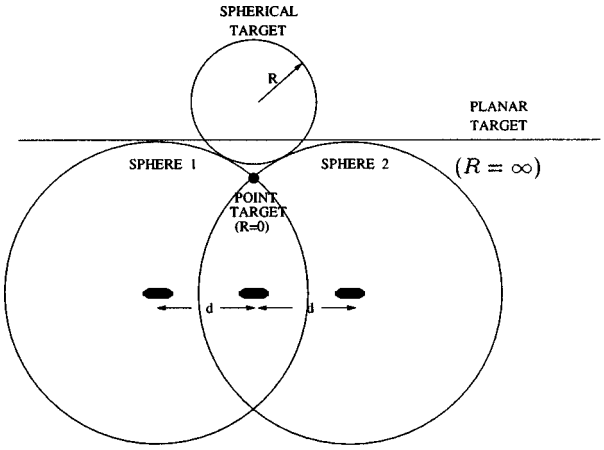


FIG. 4. The indeterminacy of curvature with only two measurements. The unknown target can have any curvature from zero to infinity.

$$\hat{\mathbf{m}}\triangleq\begin{bmatrix}\hat{z}_0\\ \hat{z}_r\\ \hat{z}_l\\ \hat{z}_u\\ \hat{z}_d\end{bmatrix},$$

$$\mathbf{z}(r,\theta,\phi,R)\triangleq\begin{bmatrix}r-R\\ \sqrt{r^2+d^2-2\,dr\cos\phi\sin\theta}-R\\ \sqrt{r^2+d^2+2\,dr\cos\phi\sin\theta}-R\\ \sqrt{r^2+d^2-2\,dr\sin\phi}-R\\ \sqrt{r^2+d^2+2\,dr\sin\phi}-R\end{bmatrix}. \tag{9}$$

The r, θ, ϕ, and R values maximizing Eq. (8) are the maximum likelihood estimates which can be found by solving $\hat{\mathbf{m}}=\mathbf{z}(\hat{r},\hat{\theta},\hat{\phi},\hat{R})$ for $\hat{R}$, $\hat{r}$, $\hat{\theta}$, and $\hat{\phi}$:

$$\hat{R}=\frac{(\hat{z}_r^2+\hat{z}_l^2)-2(\hat{z}_0^2+d^2)}{4\hat{z}_0-2(\hat{z}_r+\hat{z}_l)}=\frac{(\hat{z}_u^2+\hat{z}_d^2)-2(\hat{z}_0^2+d^2)}{4\hat{z}_0-2(\hat{z}_u+\hat{z}_d)}, \tag{10}$$

$$\hat{r}=\hat{z}_0+\hat{R}, \tag{11}$$

$$\hat{\theta}=\sin^{-1}\left[\frac{(\hat{z}_l-\hat{z}_r)(\hat{z}_l+\hat{z}_r+2\hat{R})}{\sqrt{16d^2(\hat{z}_0+\hat{R})^2-(\hat{z}_d-\hat{z}_u)^2(\hat{z}_d+\hat{z}_u+2\hat{R})^2}}\right], \tag{12}$$

$$\hat{\phi}=\sin^{-1}\left[\frac{(\hat{z}_d-\hat{z}_u)(\hat{z}_d+\hat{z}_u+2\hat{R})}{4d(\hat{z}_0+\hat{R})}\right]. \tag{13}$$

To localize the center of a target with finite radius, one needs three measurements in 2-D and five in 3-D as evident from the above expressions for the polar coordinates of the target.

B. Point target: The limit $R\to 0$

In the limit $R\to 0$, a point target is obtained. Point-target localization in 2-D has been considered in Ref. 40 and two methods of estimating the location have been presented using a linear array of transducers. The equations in 3-D derived above for finite R become simpler in the limit $R\to 0$:

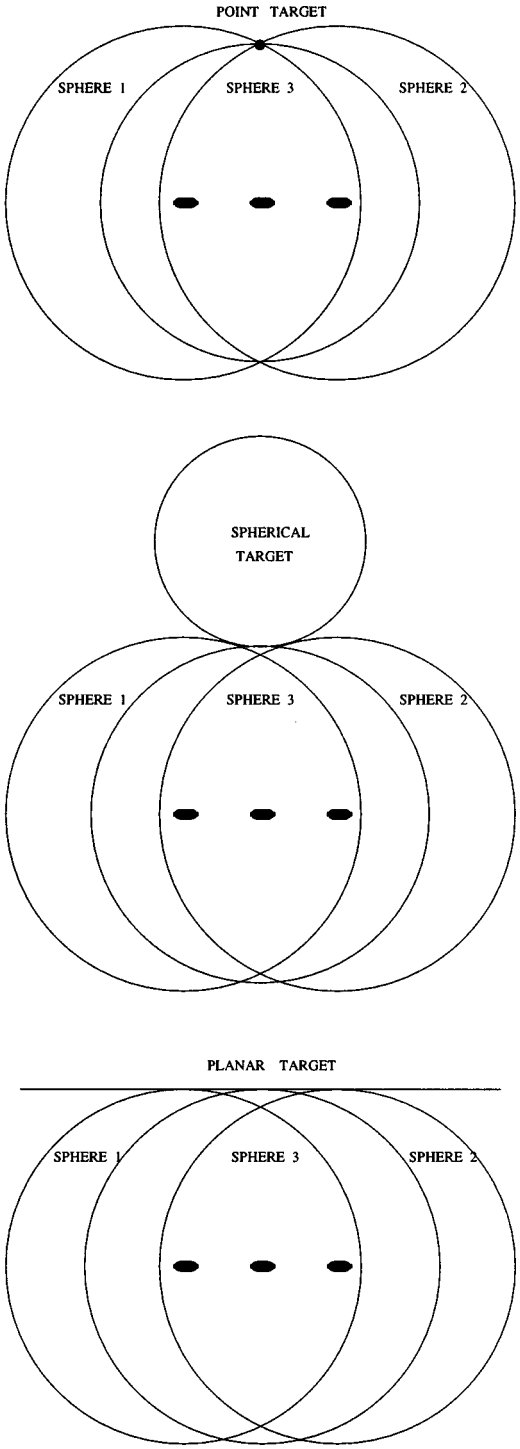


FIG. 5. Measurement geometry for point, spherical and planar targets. Three measurements uniquely identify the curvature of the unknown target both in 2-D and 3-D.

$$\hat{r}=\hat{z}_0=\sqrt{\frac{\hat{z}_r^2+\hat{z}_l^2}{2}-d^2}=\sqrt{\frac{\hat{z}_u^2+\hat{z}_d^2}{2}-d^2},$$

$$\hat{\theta}=\sin^{-1}\left[\frac{(\hat{z}_l^2-\hat{z}_r^2)}{\sqrt{16d^2\hat{z}_0^2-(\hat{z}_d^2-\hat{z}_u^2)^2}}\right] \tag{14}$$

$$\hat{\phi}=\sin^{-1}\left[\frac{(\hat{z}_d^2-\hat{z}_u^2)}{4d\hat{z}_0}\right].$$

Characterizing the point-target response of a sensor is important not only for its application to point or edgelike targets, but also to assess its performance on extended targets. There are different approaches for modeling extended targets.[12,41,42] If the approach is one of hypothesis testing or one of parameterizing the extended target, then sensor performance may not be easily related to its point-target response. On the other hand, for extended targets of unknown shape with possible roughness,[32] point-target analysis can be extremely useful.

C. Planar target: The limit $R\to\infty$

In the limit $R\to\infty$, the target becomes a plane. Both the distance to the center of the ''sphere'' and its radius of curvature become infinity. In this case, either the limits of the above equations can be taken, or more simply, the perpendicular distances of the transducers to the plane can be directly derived from the geometry. With the measurement noise taken into account:

$$\begin{aligned}
&\hat{z}_0=z_0+n_0,\\
&\hat{z}_r=z_0-d\cos\phi\sin\theta+n_r,\\
&\hat{z}_l=z_0+d\cos\phi\sin\theta+n_l,\\
&\hat{z}_u=z_0-d\sin\phi+n_u,\\
&\hat{z}_d=z_0+d\sin\phi+n_d.
\end{aligned} \tag{15}$$

The solution is:

$$\begin{aligned}
&\hat{R}=\frac{(\hat{z}_r^2+\hat{z}_l^2)-2(\hat{z}_0^2+d^2)}{4n_0-2(n_r+n_l)}\cong\infty,\\
&\hat{r}=\hat{z}_0+\hat{R}\cong\infty,\\
&\hat{\theta}=\sin^{-1}\left[\frac{\hat{z}_l-\hat{z}_r}{\sqrt{4d^2-(\hat{z}_d-\hat{z}_u)^2}}\right],\\
&\hat{\phi}=\sin^{-1}\left[\frac{\hat{z}_d-\hat{z}_u}{2d}\right].
\end{aligned} \tag{16}$$

IV. SENSITIVITY ANALYSIS OF RADIUS OF CURVATURE

A sensitivity analysis has been performed to determine how much variation would result in the radius of curvature estimate R as a result of measurement errors and variation in certain system parameters such as z_0, z_l, z_r, d, R, and θ. Since R can be estimated using either the set of measurements z_0, z_r, z_l or the set of measurements z_0, z_u, z_d, and since the equations involved are identical in form [Eq. (10)], the sensitivity of R to z_r (z_l) is the same as its sensitivity to z_u (z_d). Furthermore, since the curvature estimation equation is symmetric with respect to z_r and z_l (and the alternative

equation is symmetric with respect to z_u and z_d), the sensitivity of the curvature with respect to the individual measurements z_r, z_l, z_u, z_d will be identical. Therefore, it is sufficient to carry out the sensitivity analysis for only one of these distance measurements, which is chosen to be z_r.

Given the radius of curvature estimate in Eq. (10), perturbation is added to the variable for which sensitivity analysis is made. For example, the perturbation Δz_0 is added to z_0 and its effect on the radius of curvature estimate is calculated as follows:

$$\begin{aligned}\Delta R &= R(z_0+\Delta z_0, z_r, z_l, d) - R(z_0, z_r, z_l, d) \\ &= \frac{(z_r^2+z_l^2)-2[(z_0+\Delta z_0)^2+d^2]}{4(z_0+\Delta z_0)-2(z_r+z_l)} \\ &\quad - \frac{(z_r^2+z_l^2)-2(z_0^2+d^2)}{4z_0-2(z_r+z_l)}. \end{aligned} \tag{17}$$

In Fig. 6(a), Eq. (17) has been plotted for r between 0 and 1.5 m. The perturbation error Δz_0 on z_0 has been varied between 0 and 0.4 mm. Actually, the typical standard deviation of the range measurement error in the current system is approximately 0.18 mm as verified in Sec. VI B. A stationary spherical target with radius 7.5 cm is assumed to be present at $\theta=0°, \phi=0°$. Transducer separation is set to its minimum value of 7.5 cm corresponding to the worst resolution case. For fixed transducer separation, error in radius of curvature increases linearly with z_0 but nonlinearly with increasing r. At a target range of $r=50$ cm, an error of $\Delta z_0 = 0.18$ mm on z_0 corresponds to an error of $\Delta R = 15.4$ mm on R, which represents about 20% error. At $r=1.0$ m, the same perturbation error on z_0 corresponds to an error of $\Delta R = 61.0$ mm on curvature, representing 82% error.

The fact that a positive error Δz_0 leads to a positive error ΔR can be explained as follows: Assuming that the target has convex curvature, increasing z_0 while keeping z_r and z_l constant corresponds to an increase in the radius of curvature R, or a decrease in curvature $1/R$. For fixed Δz_0, error in R also increases with range $r=z_0+R$, since the fixed transducer separation d provides poorer resolution as the range increases. For larger ranges, as the ratio d/r decreases, the system has smaller resolution for the given separation.

In Fig. 6(b), the sensitivity of R with respect to z_r (or equivalently z_l, z_u, z_d) has been plotted for the same parameters used in generating Fig. 6(a). Note that, in this case, for a positive error Δz_r or Δz_l, the error ΔR on radius of curvature is negative. This again can be explained by the geometry of Fig. 2. A positive error on the right and left measurements, with z_0 constant, causes a reduction in the radius of curvature.

Fig. 7(a) illustrates the effect of transducer separation d on the accuracy of the radius of curvature estimate. For r between 0 and 1.5 m and $\Delta z_0 = 0.18$ mm, Eq. (17) has been plotted for transducer separations between 4.0 and 60 cm. Corresponding plot for $\Delta z_r = 0.18$ mm is presented in Fig. 8(a). In both figures, it is observed that the ratio d/r is a significant parameter in the curvature estimation process.

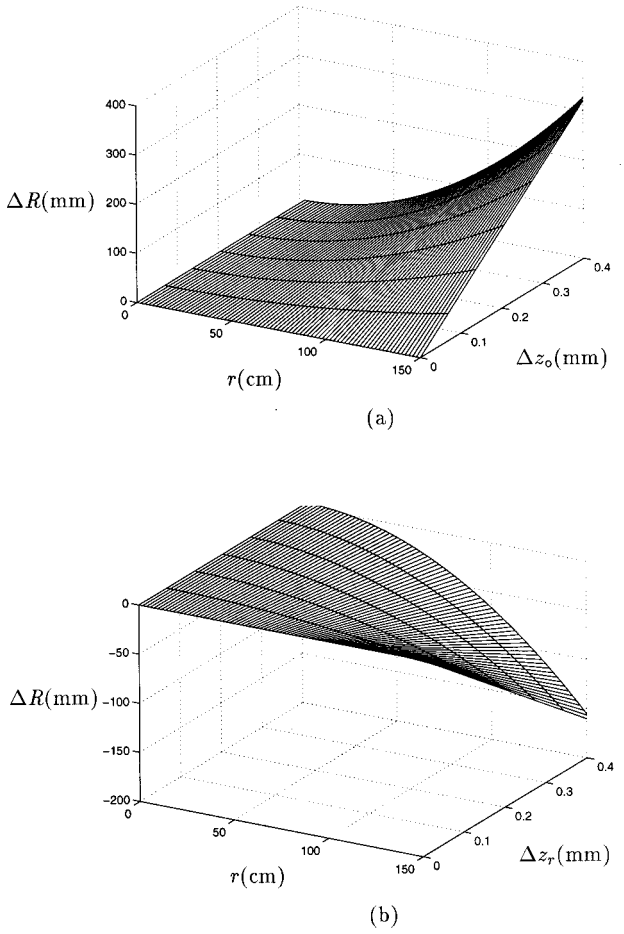


FIG. 6. Sensitivity of R to distance measurements (a) z_0; (b) z_r or z_l.

This ratio cannot be selected too large. This is because the directional sensitivity regions of the transducers are limited to a cone with half angle α_0 (in the current system, $\alpha_0 \cong 12°$, Ref. 39). If d/r is selected too large, the sensitivity regions of the transducers will not overlap at the location of the target so that targets nearer than $[(d-a)/\tan\alpha_0] + (a^2/\lambda)$ will not be detected.[43] This approximately corresponds to the distance between the central transducer and the start of the joint sensitivity region in Fig. 1. On the other hand, if d/r is selected too small, the resolution provided by the differential time-of-flight information between the central and surrounding transducers will not be sufficiently large to estimate the curvature reliably. Hence, as the operating range increases, a larger transducer separation must be used in order to maintain the same accuracy in curvature estimation. This is verified by the experimental results in Sec. VI B where it can be observed that if the range is increased while the transducer separation is kept constant, the error on R increases. Thus, it is concluded that a sensor system which is to operate over a large range of target distances must have the capability of adaptively adjusting transducer separation d. The information provided by Figs. 1 and 6(a) can be combined to formulate a rule for choosing the optimal transducer separation d for a given range r. Thus, one can envisage a two-step curvature estimation process: The range estimate obtained in the first step is used to adjust the transducer

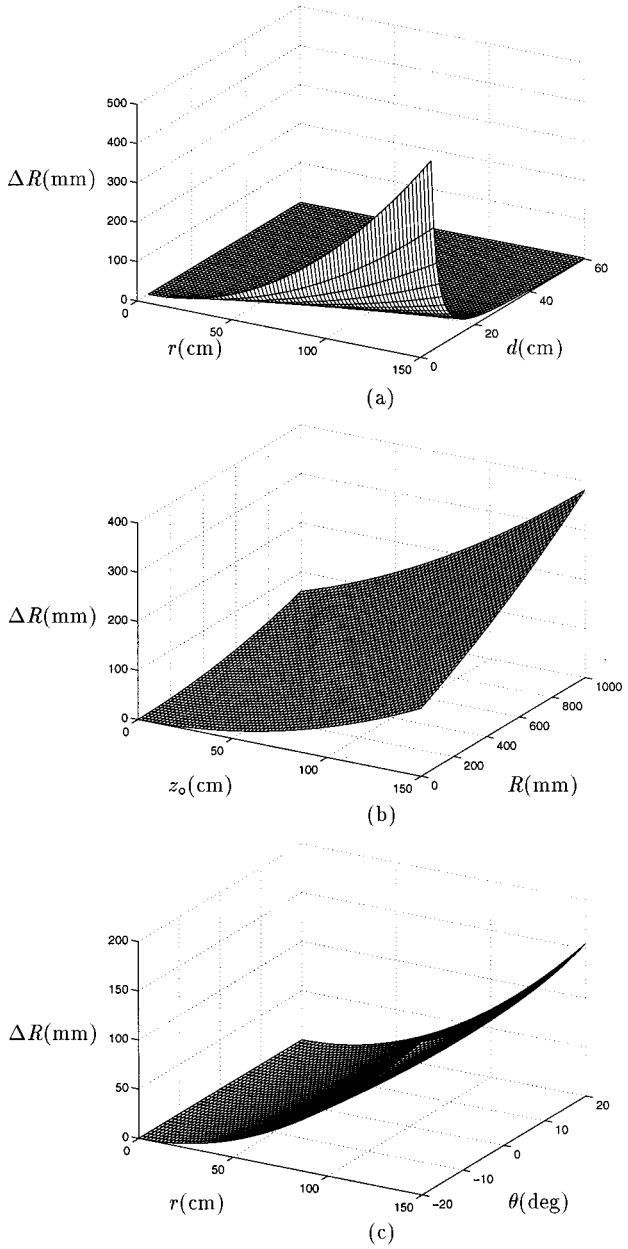


FIG. 7. Sensitivity of R to (a) d, (b) R and (c) θ when $\Delta z_0 = 0.18$ mm.

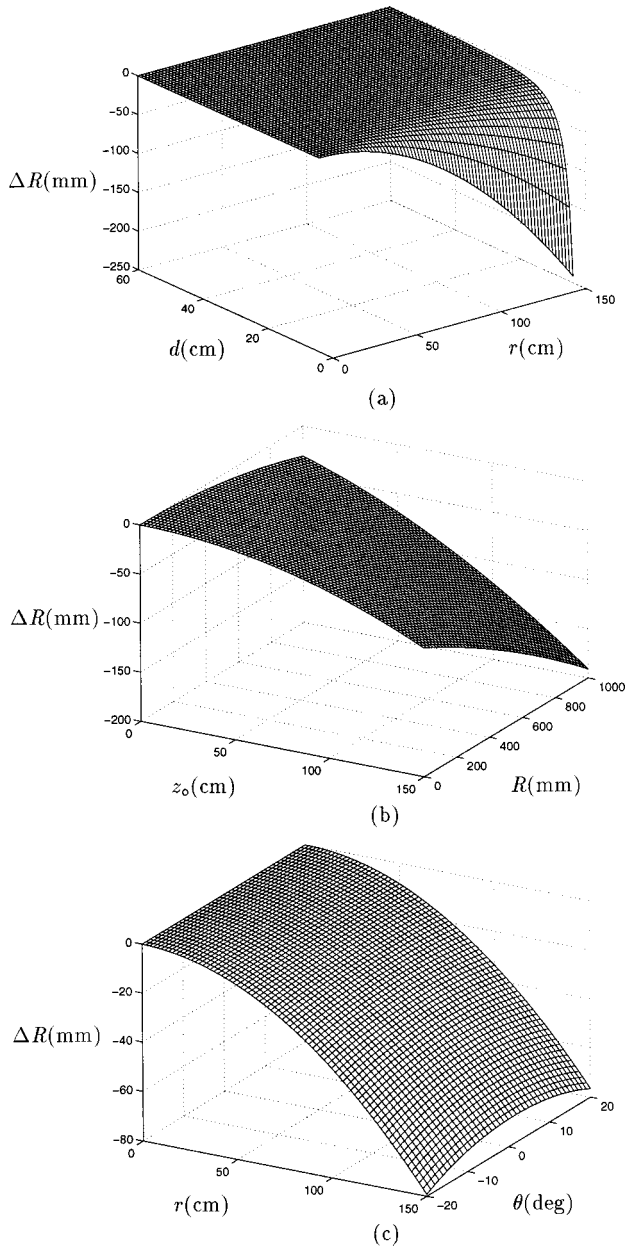


FIG. 8. Sensitivity of R to (a) d, (b) R and (c) θ when $\Delta z_r = 0.18$ mm.

separation d to its optimal value, allowing an accurate curvature estimate in the second step.

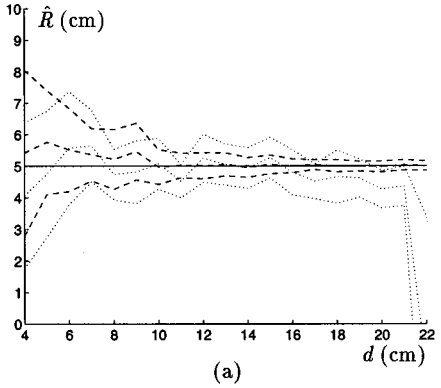
The sensitivity of curvature to a measurement error Δz_0 at different curvature and azimuth values has also been investigated when $\Delta z_0 = 0.18$ mm. In the first case, it is observed that for fixed z_0, ΔR increases with the radius of curvature of the target as illustrated in Fig. 7(b). In Fig. 7(c), θ is varied from $-20°$ to $20°$ with $d = 7.5$ cm. It can be observed that there is a slight increase in ΔR with increasing $|\theta|$. Corresponding plots for a perturbation $\Delta z_r = 0.18$ mm on the radius of curvature estimate are presented in Fig. 8(b) and (c).

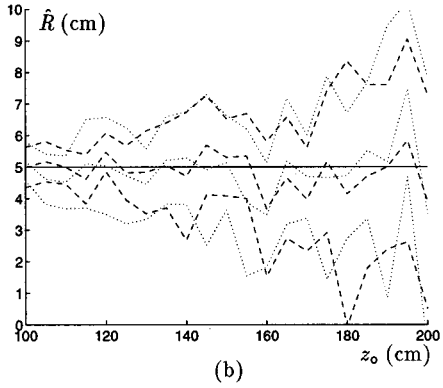
As will also be confirmed experimentally in Sec. VI B, accuracy is much less of a problem with the localization parameters so that a sensitivity analysis is not presented for these parameters.

V. SIMULATION RESULTS

In the simulations, the signals are modeled according to Eq. (4) using the parameter values $A_{\max} = 1$, $r_{\min} = 5.8$ cm, $\rho_c = 0.45R - 0.022$, $f_0 = 49.4$ kHz, $c = 343.3$ m/s. Zero-mean white Gaussian noise $N(0, \sigma)$ of varying magnitude has been added to the signals before estimating the time-of-flight.

The initial estimate of the radius of curvature is made using Eq. (10) with the first set of measurements obtained at the initial (flat) position of the sensor (Fig. 2). The target position (r, θ, ϕ) is estimated using Eqs. (11)–(13). Using these initial estimates, transducers will be rotated toward the object so that it is located along their line-of-sights in order to obtain more nearly accurate results. For this purpose, the inclination angles θ_r, θ_l of the right and left transducers with respect to the target are estimated from the initial noisy measurements:





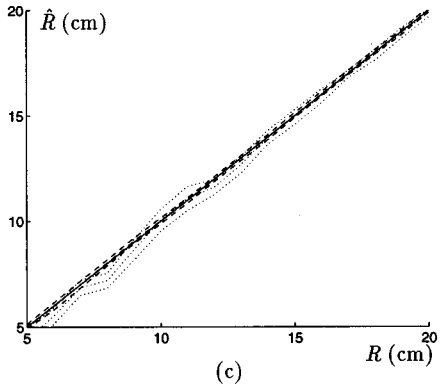
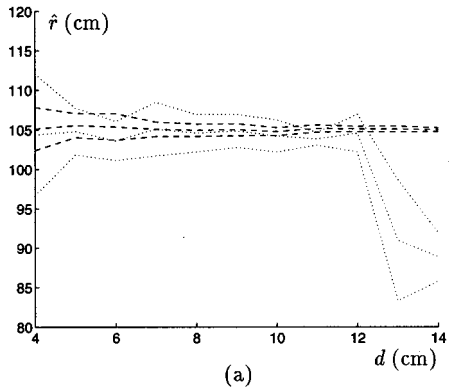
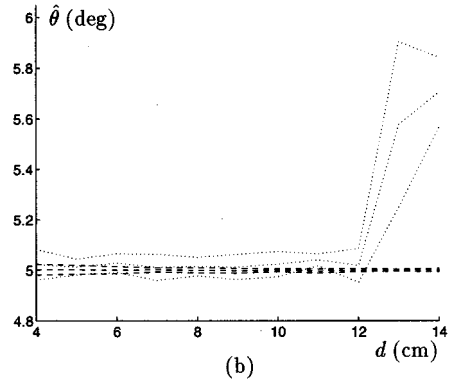


FIG. 9. Estimated radius versus (a) d, (b) z_0, and (c) R. Dotted and dashed lines indicate the mean of the estimate and $\pm\sigma_{\hat{R}}$ obtained at the flat and adapted positions, respectively. In part (a), $R=5$ cm, $z_0=100$ cm. In part (b), $R=5$ cm, $d=10$ cm. In part (c), $d=20$ cm, $z_0=100$ cm.





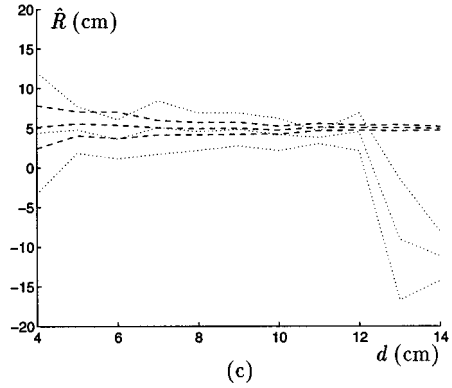


FIG. 10. Location and radius of curvature estimation versus d. Dotted and dashed lines indicate the mean of the estimate and $\pm\sigma_{\hat{R}}$ obtained at the flat and adapted positions respectively. In all parts of the figure, $R=5$ cm, $z_0=100$ cm, and $\theta=5°$.

$$\hat{\theta}_r=\tan^{-1}\left(\frac{\hat{r}\cos\hat{\phi}\sin\hat{\theta}}{\hat{r}\cos\hat{\phi}\sin\hat{\theta}-d}\right),$$
$$\hat{\theta}_l=\tan^{-1}\left(\frac{\hat{r}\cos\hat{\phi}\sin\hat{\theta}}{\hat{r}\cos\hat{\phi}\sin\hat{\theta}+d}\right). \tag{18}$$

First, the central, right and left transducers are all rotated by $\hat{\phi}$ in elevation, and then rotated by $\hat{\theta}$, $\hat{\theta}_r$, and $\hat{\theta}_l$, respectively, so that they are approximately perpendicular to the object surface. New measurements are made and the radius of curvature is estimated for the second time, again using Eq. (10). Flat and rotated configurations of the transducers are illustrated in Fig. 2 in 2-D. The distance measurements z_0, z_r, z_l, z_u, z_d should ideally be the same before and after the rotation. In practice, they will be slightly different due to the additive measurement noise. In the second estimate, the distance measurements z_0, z_r, z_l, z_u, z_d are obtained with a larger SNR since the transducers are now looking straight ahead at the object. This in turn results in more nearly accurate estimates.

In all of the simulations, a 100-realization Monte-Carlo study has been employed. Figure 9(a) illustrates the effect of varying d on the curvature estimate before and after the transducers are rotated. Note the reduction in the standard deviations of the estimates as d is increased. For $d>21$ cm, all of the transducers cannot detect the object and the estimates become very erroneous. Figure 9(b) illustrates that as z_0 increases, standard deviations of both estimates increase.

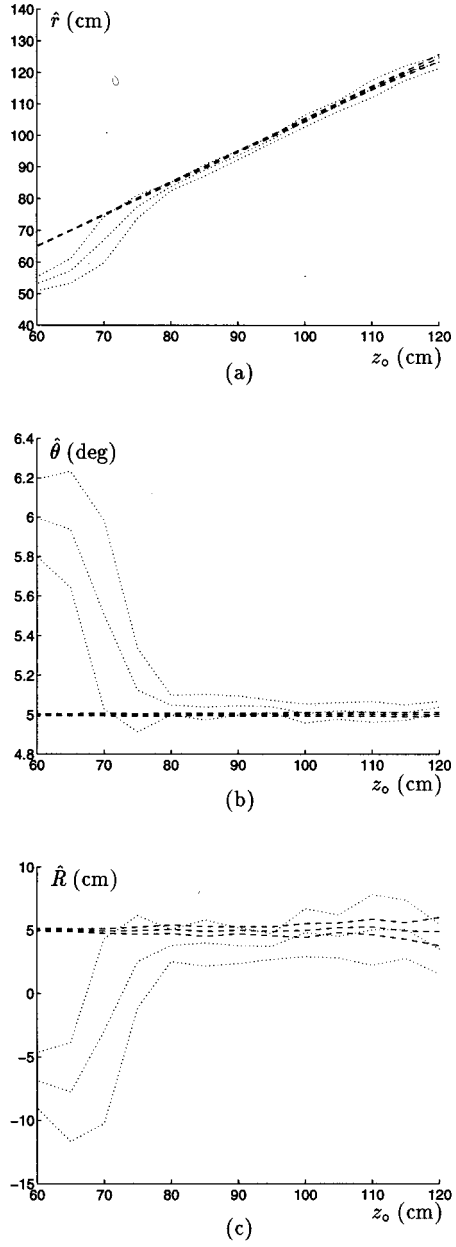


FIG. 11. Location and radius of curvature estimation versus z_0. Dotted and dashed lines indicate the mean of the estimate and $\pm\sigma_{\hat{R}}$ obtained at the flat and adapted positions, respectively. In all parts of the figure, $R=5$ cm, $d=10$ cm, $z_0=100$ cm, and $\theta=5°$.

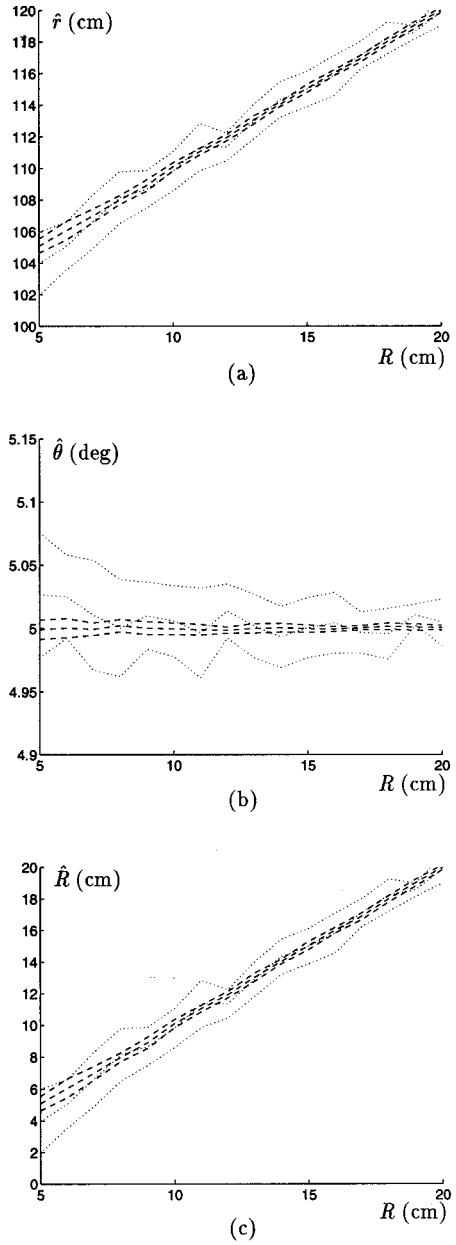
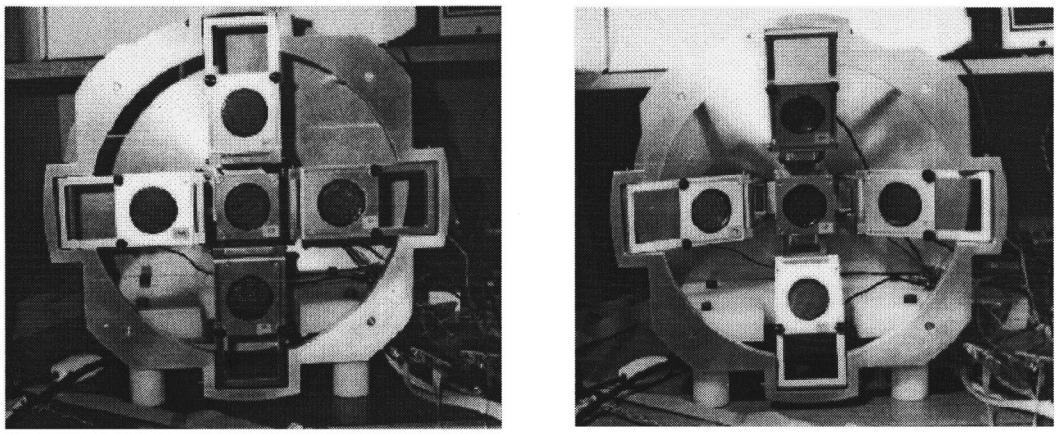


FIG. 12. Location and radius of curvature estimation versus R. Dotted and dashed lines indicate the mean of the estimate and $\pm\sigma_{\hat{R}}$ obtained at the flat and adapted positions, respectively. In all parts of the figure, $d=10$ cm, $z_0=100$ cm, and $\theta=5°$.

Figure 9(c) illustrates the effect of the true radius. In all cases, significant improvement can be observed after the transducers are adapted.

Figure 10 shows the dependence of the estimates on d for $R=5$ cm, $z_0=100$ cm, and $\theta=5°$. Figure 10(a) illustrates that as d increases, the estimate after rotation keeps improving. The estimation before rotation improves up to $d=12$ cm, and after that point, it gets worse since the target is now located either at very low-SNR regions of the sensitivity region or outside it (as d increases, the normal angles of the left and right sensors increase). The average error in $\hat{R}$ before rotation is 5.96%. After rotation, the error is reduced to 0.05%. Figure 10(b) illustrates how the azimuth estimate $\hat{\theta}$ is affected by d. The average error figures before and after rotation are 6.0% and 0.0% respectively. Figure 10(c) illustrates the dependence of the curvature estimate on d. The average error before and after rotation is 11.4% and 1.2%, respectively.

Figure 11 shows the dependence of $\hat{r}, \hat{\theta}$, and $\hat{R}$ estimates on the distance z_0 for $R=5$ cm, $d=10$ cm, $z_0=100$ cm, and $\theta=5°$. For small values of z_0, the normal angles of the left and right sensors are large and the initial estimates obtained with the flat configuration are not very accurate. For large values of z_0, the normal angles decrease and the estimation accuracy improves. In Fig. 11(a), the average error of the distance estimate is 6.0% before rotation and 0.0% after rotation. In Fig. 11(c), the average error of radius of curvature estimation before rotation is 19.0% and 0.0% after rotation.



(a) (b)

FIG. 13. Two extreme positions of the sensing device.

Again, there is a minimum value of $z_0 \cong 70$ cm, at which the target is accurately detectable. For shorter distances, all estimates become erroneous since one or more transducers cannot detect the target.

Figure 12 shows the effect of varying R on the $\hat{r}, \hat{\theta}$, and $\hat{R}$ estimates for $d=10$ cm, $z_0=100$ cm, and $\theta=5°$. As R increases, the accuracy of all three estimates improve. In addition, there is significant improvement after the transducers are adapted to the target, as evident from the reduction in the standard deviations of the estimates.

VI. EXPERIMENTAL VERIFICATION

A. The sensing device and the experimental setup

The sensing device used in this investigation was precision constructed for 3-D sonar applications. A trade-off between simplicity of mechanical design and flexibility was established. The unit, illustrated in Fig. 13, consists of five Polaroid 6500 series acoustic transducers, each operating at a resonance frequency of $f_0=49.4$ kHz.[39] A central transducer is flanked by four transducers symmetrically. The position of the central transducer is fixed but the separation d of each surrounding transducer from the center can be manually adjusted between 7.5 to 12.0 cm.

The device has 16 mechanical joints which enable it to move with the aid of stepper motors located behind the central transducer. In one extreme position, all the sensors are coplanar [Fig. 13(a)]. In the other extreme, the flanking transducers have been rotated by 30° as shown in Fig. 13(b). In between, there are 1020 intermediate positions. The flexibility of the sensor can be judiciously used to recognize 3-D targets and focus on them to use the sensor data more effectively.

The analysis of the previous sections has been verified by real sonar data from cylindrical and planar targets using a 4-channel DAS-50 A/D card with 12-bit resolution and 1 MHz sampling frequency. Echo signals were processed on an IBM-PC 486 using the C programming language. The block diagram for the hardware is shown in Fig. 14. The experiments were conducted in 2-D to allow accurate calibration. Real distances were ascertained accurately by carrying out the whole set of experiments on large sheets of millimetric paper. Transducer separation was kept constant at $d=7.5$ cm (except in the results presented in Table III where $d=12.0$ cm and those in Table VII where d is varied).

Each transducer was made to transmit and receive in sequence to avoid crosstalk and to benefit maximally from the high sampling rate of the A/D card. Starting at the transmit time, 10 000 samples of each echo signal have been collected to estimate the time-of-flight. The targets employed in this study are: cylinders with radii 25 mm, 50 mm, 75 mm and a planar target. All targets used in the experiments were wooden, with smooth surfaces, each with a height of 120 cm. In the experiments, the sensors were situated at the middle of this height so that the lower and upper edges of the cylinders remained outside the beam patterns of the transducers. For the maximum range considered in the experiments (150 cm), from the geometry, the beam extends over a transverse extent of $150\times 2\times \tan\alpha_0 = 64$ cm, which is less than 120 cm. Since the edges remain outside the beam patterns, edge effects are not observed. The cylindrical target with radius 25 mm is considered a good approximation to an edge target since cylinders with small radii behave similarly to outer edges formed by the intersection of two planes.[44]

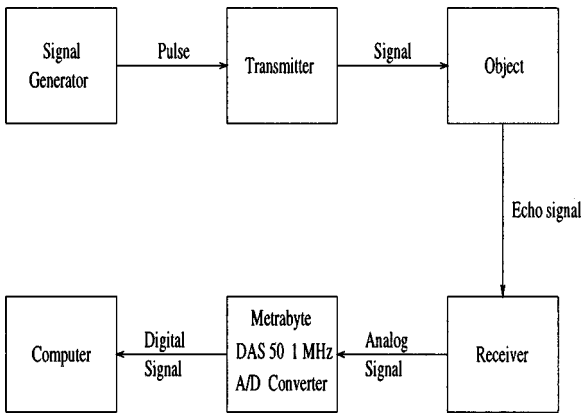


FIG. 14. The block diagram of the experimental setup.

TABLE I. Experimental results for a cylinder of $R=75$ mm when $d=7.5$ cm.

z_0 (mm)	$E\{\hat{z}_0\}$ (mm)	$\sigma_{\hat{z}_0}$ (mm)	$E\{\hat{\theta}\}$ (deg)	$\sigma_{\hat{\theta}}$ (deg)	$E\{\hat{R}_1\}$ (mm)	$\sigma_{\hat{R}_1}$ (mm)	$E\{\hat{R}_2\}$ (mm)	$\sigma_{\hat{R}_2}$ (mm)
300	299.93	0.14	0.01	0.08	73.39	7.78	74.47	4.67
400	400.00	0.14	0.02	0.07	78.83	11.77	76.28	6.82
500	500.15	0.14	−0.04	0.07	75.14	17.13	75.16	9.76
600	600.02	0.14	−0.04	0.07	75.65	24.39	75.99	14.14
700	700.10	0.18	0.01	0.07	81.72	32.00	78.87	18.24
800	800.15	0.14	0.02	0.07	80.30	38.76	77.14	22.10
900	900.42	0.14	0.07	0.07	71.12	50.33	73.62	28.18
1000	1000.24	0.16	−0.02	0.07	76.25	60.24	75.39	34.33
1100	1100.23	0.14	−0.03	0.07	75.55	70.94	76.24	39.02
1200	1200.07	0.14	0.02	0.07	80.50	84.47	79.41	46.46
1300	1300.19	0.14	0.00	0.07	78.71	95.51	76.33	53.49
1400	1400.37	0.14	0.01	0.07	80.08	124.35	78.85	68.40
1500	1500.22	0.15	0.01	0.07	69.19	118.16	75.39	66.17

TABLE III. Experimental results for a cylinder of $R=50$ mm with $d=12.0$ cm.

z_0 (mm)	$E\{\hat{z}_0\}$ (mm)	$\sigma_{\hat{z}_0}$ (mm)	$E\{\hat{\theta}\}$ (deg)	$\sigma_{\hat{\theta}}$ (deg)	$E\{\hat{R}_1\}$ (mm)	$\sigma_{\hat{R}_1}$ (mm)	$E\{\hat{R}_2\}$ (mm)	$\sigma_{\hat{R}_2}$ (mm)
300	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
400	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
500	499.75	0.14	−0.07	0.05	48.24	6.21	48.50	3.75
600	599.97	0.17	−0.08	0.05	49.54	8.96	49.08	5.47
700	700.10	0.15	0.00	0.05	47.20	11.43	48.16	7.20
800	799.93	0.15	0.02	0.05	48.58	15.21	48.01	9.13
900	900.20	0.16	0.00	0.05	52.20	19.08	50.33	11.07
1000	999.77	0.15	−0.02	0.05	46.84	22.51	48.84	13.28
1100	1100.43	0.16	−0.03	0.05	55.43	27.45	53.69	15.37
1200	1200.11	0.16	−0.03	0.05	45.14	31.74	47.83	18.41
1300	1300.34	0.15	−0.02	0.04	48.48	38.63	48.50	21.63
1400	1400.26	0.17	−0.02	0.05	47.48	43.53	49.06	25.81
1500	1500.37	0.14	−0.07	0.04	49.37	46.80	50.29	25.74

B. Experimental results

Each target's surface distance z_0 to the central transducer was varied between 30 and 150 cm at 10 cm intervals. At each distance, data were collected while the target was stationary at $\theta=0°$. For the same target position, 1000 sets of measurements were taken. Each set of measurements provides a single estimate of target radius of curvature, range and azimuth. The typical differential distance measured by the central and the right/left transducers varies between 0 and 10 mm depending on the target curvature and distance, for $d=7.5$ cm. As the range of the target increases, the differential becomes less reliable to extract the curvature information.

The means and standard deviations of $\hat{z}_0$, $\hat{\theta}$, and $\hat{R}$ of each type of target considered are computed and tabulated in Tables I–V. In all of the tables, results for $\hat{z}_0$ and $\hat{R}$ are tabulated individually instead of r itself, which is the sum of these two components. Radius of curvature estimates have been presented both at the flat position and adapted position of the transducers. The results before and after adaptation are denoted by the subscripts 1 and 2, respectively. The results for $\hat{z}_0$ and $\hat{\theta}$ do not vary much after the rotation. Therefore, for these parameters, only the initial estimates made at the flat position have been presented.

In Table III, results for $z_0=300$ mm and $z_0=400$ mm have not been presented since a very thin cylinder at $\theta=0°$ is not detectable by the right and left transducers at the transducer separation of $d=12.0$ cm. For the same reason, results for $z_0=300$ mm in Table IV are excluded when d was set equal to 7.5 cm. From the results, it can be observed that the $\hat{z}_0$ and $\hat{\theta}$ estimates are quite accurate: For a stationary target, the typical standard deviation of $\hat{z}_0$ is 0.18 mm, which is approximately the resolution allowed by the A/D converter sampling rate. The typical standard deviation of the azimuth estimate is 0.08°. Error on $\hat{z}_0$ and $\hat{\theta}$ are relatively constant as the distance of the target is varied between 30 and 150 cm. However, for the curvature, typical error is around 7–9 mm at 30 cm, but keeps increasing with range for a fixed transducer separation. This is due to the reduction in the d/r ratio which provides poorer resolution in estimating curvature. To estimate the curvature of a cylindrical target reliably, it is necessary to increase the transducer separation as the range is increased as seen in Fig. 6(a). To illustrate the effect of transducer separation, results for the maximum allowed separation in the system ($d=12$ cm) are included in

TABLE II. Experimental results for a cylinder of $R=50$ mm when $d=7.5$ cm.

z_0 (mm)	$E\{\hat{z}_0\}$ (mm)	$\sigma_{\hat{z}_0}$ (mm)	$E\{\hat{\theta}\}$ (deg)	$\sigma_{\hat{\theta}}$ (deg)	$E\{\hat{R}_1\}$ (mm)	$\sigma_{\hat{R}_1}$ (mm)	$E\{\hat{R}_2\}$ (mm)	$\sigma_{\hat{R}_2}$ (mm)
300	300.00	0.14	−0.03	0.09	48.27	6.84	48.93	4.10
400	399.79	0.14	−0.02	0.07	51.13	10.08	50.27	6.25
500	500.05	0.16	−0.04	0.07	49.17	15.23	49.44	9.29
600	600.59	0.17	0.04	0.07	52.33	22.43	51.79	13.24
700	699.98	0.15	0.01	0.07	44.96	27.61	46.63	16.57
800	799.87	0.14	−0.03	0.07	44.55	35.02	48.36	21.52
900	900.40	0.15	0.05	0.07	50.64	45.42	50.25	26.80
1000	1000.39	0.14	0.03	0.07	59.87	57.41	55.48	33.39
1100	1100.17	0.15	−0.10	0.07	53.83	64.65	52.90	38.14
1200	1200.14	0.14	0.01	0.07	48.08	83.71	49.21	47.71
1300	1300.15	0.14	−0.05	0.07	63.70	91.43	58.88	52.12
1400	1400.63	0.15	−0.08	0.07	40.92	122.53	45.62	69.84
1500	1500.18	0.15	0.01	0.07	46.55	125.46	49.31	74.02

TABLE IV. Experimental results for a cylinder of $R=25$ mm when $d=7.5$ cm.

z_0 (mm)	$E\{\hat{z}_0\}$ (mm)	$\sigma_{\hat{z}_0}$ (mm)	$E\{\hat{\theta}\}$ (deg)	$\sigma_{\hat{\theta}}$ (deg)	$E\{\hat{R}_1\}$ (mm)	$\sigma_{\hat{R}_1}$ (mm)	$E\{\hat{R}_2\}$ (mm)	$\sigma_{\hat{R}_2}$ (mm)
300	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
400	400.06	0.16	0.02	0.08	21.07	9.14	23.83	5.02
500	500.11	0.14	−0.03	0.07	23.58	14.00	22.49	7.84
600	599.99	0.15	0.04	0.07	28.58	20.07	27.25	10.82
700	699.99	0.15	−0.05	0.07	26.20	26.01	24.67	13.53
800	799.77	0.16	0.02	0.07	26.51	34.23	25.50	19.17
900	899.91	0.20	0.02	0.07	30.37	42.98	27.94	22.78
1000	999.44	0.24	0.03	0.08	29.86	57.48	26.90	34.49
1100	1100.18	0.17	0.02	0.07	28.52	64.30	24.32	35.37
1200	1200.44	0.16	−0.04	0.07	23.47	78.03	27.65	41.37
1300	1300.31	0.18	−0.04	0.07	23.67	91.49	23.80	51.95
1400	1400.01	0.15	0.00	0.07	23.50	98.58	24.37	55.01
1500	1499.39	0.25	0.03	0.08	23.01	139.20	26.22	76.26

TABLE V. Experimental results for a planar target of $R=\infty$ when $d=7.5$ cm.

z_0 (mm)	$E\{\hat{z}_0\}$ (mm)	$\sigma_{\hat{z}_0}$ (mm)	$E\{\hat{\theta}\}$ (deg)	$\sigma_{\hat{\theta}}$ (deg)	$E\{\hat{R}_1\}=E\{\hat{R}_2\}$ (mm)	$\sigma_{\hat{R}_1}=\sigma_{\hat{R}_2}$ (mm)
300	299.88	0.14	−0.02	0.08	4.70×10^{18}	2.12×10^{19}
400	399.94	0.14	0.01	0.08	1.30×10^{19}	3.36×10^{19}
500	500.09	0.15	−0.04	0.08	8.00×10^{18}	2.72×10^{19}
600	600.11	0.16	−0.01	0.07	1.42×10^{19}	3.49×10^{19}
700	699.98	0.16	−0.02	0.08	1.53×10^{19}	3.60×10^{19}
800	800.24	0.15	−0.06	0.08	8.00×10^{18}	2.71×10^{19}
900	899.83	0.17	0.03	0.08	1.08×10^{19}	3.10×10^{19}
1000	1000.07	0.15	0.03	0.08	1.25×10^{19}	3.31×10^{19}
1100	1100.18	0.16	0.04	0.08	1.23×10^{19}	3.28×10^{19}
1200	1199.92	0.17	−0.04	0.08	7.90×10^{18}	2.70×10^{19}
1300	1300.00	0.18	−0.05	0.08	9.20×10^{18}	2.89×10^{19}
1400	1399.80	0.15	−0.02	0.08	1.36×10^{19}	3.43×10^{19}
1500	1500.34	0.20	0.00	0.08	1.07×10^{19}	3.09×10^{19}

Table III. Compared to Table II where $d=7.5$ cm, it is observed that errors in the radius of curvature estimate are approximately reduced by 60%. In Table V, results for a planar target ($R=\infty$) are illustrated. In the experiments, whenever the denominator of Eq. (10) is zero, a very large value (10^{20}) is assigned to R to be able to represent it numerically.

In Table VI, results for the cylinder with $R=25$ mm are provided for $\theta=0°,3°,5°,8°$. It is observed that the accuracies of range and azimuth estimates do not change significantly as compared to the case when the target is along the line-of-sight. The accuracy of the initial curvature estimate degrades with $|\theta|$ as expected. However, the estimates with the adapted configuration for target at different θ are comparable in accuracy. For larger values of θ than considered in the table, it is not possible to estimate the curvature since the target will be outside the sensitivity region of either the right or the left transducer.

Finally, the transducers were detached from the mounting and were placed on polyamid stands so that larger transducer separations than allowed by the prototype system could be tested (d: 15.0–30.0 cm). The results are presented in Table VII. When $d>21$ cm, it is not possible to acquire data with the side transducers at the flat position since the target at $z_0=1.00$ m remains outside the joint sensitivity region of the transducers. Therefore, for these cases, the transducers are maintained approximately perpendicular to the object surface while experimental data are being collected.

Overall, the results indicate that the accuracy of the curvature estimation after adapting the transducers brings an improvement varying between 35% and 45%.

TABLE VI. Experimental results for a cylinder with $R=25$ mm for varying θ when $d=7.5$ cm.

θ (deg)	$E\{\hat{z}_0\}$ (mm)	$\sigma_{\hat{z}_0}$ (mm)	$E\{\hat{\theta}\}$ (deg)	$\sigma_{\hat{\theta}}$ (deg)	$E\{\hat{R}_1\}$ (mm)	$\sigma_{\hat{R}_1}$ (mm)	$E\{\hat{R}_2\}$ (mm)	$\sigma_{\hat{R}_2}$ (mm)
0	1000.26	0.16	−0.04	0.07	22.65	56.81	24.64	30.60
3	999.46	0.15	2.73	0.08	24.57	58.50	26.77	31.92
5	1000.67	0.15	4.98	0.07	25.64	61.16	23.43	29.75
8	1000.05	0.16	7.53	0.07	27.34	63.34	25.06	31.42

TABLE VII. Experimental results for varying d for a cylinder of radius $R=75$ mm, located at $z_0=1.00$ m and $\theta=0°$.

d (cm)	$E\{\hat{z}_0\}$ (mm)	$\sigma_{\hat{z}_0}$ (mm)	$E\{\hat{\theta}\}$ (deg)	$\sigma_{\hat{\theta}}$ (deg)	$E\{\hat{R}_1\}$ (mm)	$\sigma_{\hat{R}_1}$ (mm)	$E\{\hat{R}_2\}$ (mm)	$\sigma_{\hat{R}_2}$ (mm)
15	999.65	0.14	0.01	0.07	73.61	20.49	76.29	12.50
20	1000.57	0.14	−0.04	0.07	72.65	18.87	73.46	10.96
25	999.48	0.16	0.02	0.08	⋯	⋯	77.61	9.75
30	999.88	0.15	0.03	0.07	⋯	⋯	72.28	8.31
35	999.26	0.13	0.05	0.07	⋯	⋯	75.73	6.24
40	1000.74	0.15	0.04	0.06	⋯	⋯	74.05	5.49

VII. DISCUSSION AND CONCLUSION

A sensing device capable of estimating the location and radius of curvature of spherical and cylindrical targets has been described. The main goal of the study is to assess the performance of radius of curvature estimation. The estimation accuracy can be improved by employing an adaptive sensor configuration: After acquiring the initial data, transducers are rotated to align their line-of-sights with the object. This way, SNR is increased and more nearly accurate estimates can be obtained. Two limiting cases are of special interest: the point (in 3-D) or line (in 2-D) target and the planar target. Analytical results are verified by real sonar data from cylindrical and planar targets. Typical accuracies in range and azimuth are 0.18 mm and 0.1°, respectively. Accuracy of the curvature estimate depends on the target type and system parameters such as transducer separation and operating range. The estimation with the adapted configuration gives much better results than without adaptation, and brings an improvement in the accuracy of about 40%.

The radius of curvature estimation provides valuable information for differentiating reflectors with different radii (including $R=0$ for an edgelike reflector to $R=\infty$ for a planar reflector). The classification procedure, consistent with the experimental results, is illustrated in Fig. 15. The uncertainty region of each radius estimate is considered to be between $[\hat{R}-3\sigma_{\hat{R}},\hat{R}+3\sigma_{\hat{R}}]$ assuming zero-mean Gaussian-distributed estimation error. The standard deviation σ_R increases with the radius of curvature. Given two targets with constant curvature, if there is overlap between their uncertainty regions, then these targets may not be distinguished for estimates which fall within the overlap region, shown by

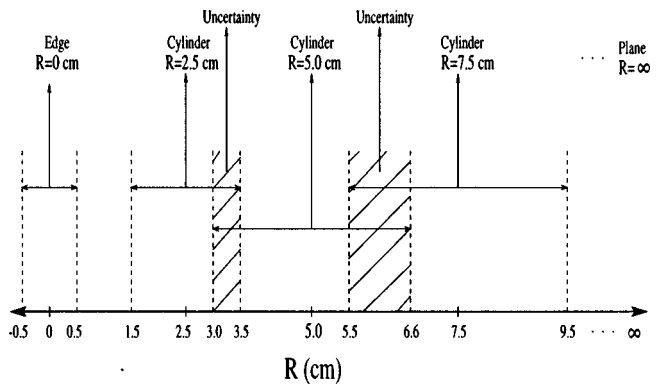


FIG. 15. Target discrimination using radius of curvature estimation.

hatched areas in Fig. 15. The results presented allow such decisions to be based on a solid footing.

For reliable curvature estimation, it is necessary to increase the transducer separation as the range is increased. The transducer separation in the present system is relatively limited and not capable of real-time dynamic adaptation. A system which is adaptive also in this respect would be able to maintain high accuracy over a broader range of distances.

When dealing with shapes more general than cylinders and spheres, such as ellipsoidal surfaces, the geometry will be slightly more complicated. Nevertheless, a similar approach can be taken, possibly requiring additional sensors, or sensors with greater capability of motion. Finally, although the method has been developed for convex ($R>0$) reflectors, it is equally applicable to concave ($R<0$) reflectors. Targets that have spatially varying curvature which may become both concave and convex have also been addressed in recent work.[45,46] Naturally, the larger the number of parameters (or degrees of freedom) of the surface, the larger the number of sensors needed.

ACKNOWLEDGMENTS

This work was supported by TÜBİTAK under Project No. EEEAG-92 and the British Council Academic Link Program. The authors would like to thank the anonymous reviewers for their comments.

[1] G. C. Carter and E. R. Robinson, ''Ocean effects on time-delay estimation requiring adaptation,'' IEEE J. Ocean Eng. **18**, 367–378 (1993).
[2] A. C. Smith and G. C. L. Searle, ''Empirical observations of a sonar adaptive array,'' IEE Proc. F, Commun. Radar Signal Process. **132**, 595–597 (1985).
[3] K. Scarbrough, G. C. Carter, and R. J. Tremblay, ''Performance predictions for coherent and incoherent processing techniques of time-delay estimation,'' IEEE Trans. Acoust., Speech, Signal Process. **31**, 1191–1196 (1983).
[4] G. C. Carter, ''Coherence and time-delay estimation,'' Proc. IEEE **75**, 236–255 (1987).
[5] M. Wazenski and D. Alexandrou, ''Active, wideband detection and localization in an uncertain multipath environment,'' J. Acoust. Soc. Am. **101**, 1961–1970 (1997).
[6] D. J. W. Hardie and A. B. Gallaher, ''Review of numerical methods for predicting sonar array performance,'' IEE Proc. F, Radar Sonar and Navigation **143**, 196–203 (1996).
[7] R. Smith, A. Stevens, A. Frost, and P. Probert, ''Developing a sensor-based underwater navigation system,'' Int. J. Syst. Sci. **29**, 1145–1155 (1998).
[8] W. K. Stewart, ''3-dimensional stochastic modeling using sonar sensing for undersea robotics,'' Autonomous Robots **3**, 121–143 (1996).
[9] R. Kuc, ''Three-dimensional tracking using qualitative bionic sonar,'' Robotics Autonomous Syst. **11**, 213–219 (1993).
[10] M. L. Hong and L. Kleeman, ''Ultrasonic classification and location of 3-D room features using maximum likelihood estimation II,'' Robotica **15**, 645–652 (1997).
[11] L. Kleeman and H. Akbarally, ''A sonar sensor for accurate 3-D target localization and classification,'' in *Proceedings IEEE International Conference on Robotics and Automation, Nagoya, Japan, May 21–27* (IEEE, Piscataway, NJ, 1995), pp. 3003–3008.
[12] H. Peremans, K. Audenaert, and J. M. Van Campenhout, ''A high-resolution sensor based on tri-aural perception,'' IEEE Trans. Rob. Autom. **9**, 36–48 (1993).
[13] A. M. Sabatini, ''Statistical estimation algorithms for ultrasonic detection of surface features,'' in *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems, Munich, Germany, September 12–16* (IEEE, Piscataway, NJ, 1994), pp. 1845–1852.
[14] A. M. Sabatini, ''Sampled baseband correlators for in-air ultrasonic rangefinders,'' IEEE Trans. Ind. Electron. **45**, 341–350 (1998).
[15] M. K. Brown, ''The extraction of curved surface features with generic range sensors,'' Int. J. Robotics Res. **5**, 3–18 (1986).
[16] R. Kuc, ''Biologically motivated adaptive sonar system,'' J. Acoust. Soc. Am. **100**, 1849–1854 (1996).
[17] R. Kuc, ''Biomimetic sonar recognizes objects using binaural information,'' J. Acoust. Soc. Am. **102**, 689–696 (1997).
[18] A. Hilton, J. Illingworth, and T. Windeatt, ''Statistics of surface curvature estimates,'' Pattern Recogn. **28**, 1201–1221 (1995).
[19] I. E. Dror, M. Zagaeski, and C. F. Moss, ''3-dimensional target recognition via sonar—a neural network model,'' Neural Networks **8**, 149–160 (1995).
[20] J. A. Simmons, P. A. Saillant, J. M. Wotton, T. Haresign, M. J. Ferragamo, and C. F. Moss, ''Composition of biosonar images for target recognition by echolocating bats,'' Neural Networks **8**, 1239–1261 (1995).
[21] H. L. Roitblat, W. W. L. Au, P. E. Nachtigall, R. Shizumura, and G. Moons, ''Sonar recognition of targets embedded in sediment,'' Neural Networks **8**, 1263–1273 (1995).
[22] R. P. Gorman and T. J. Sejnowski, ''Learned classification of sonar targets using a massively parallel network,'' IEEE Trans. Acoust., Speech, Signal Process. **36**, 1135–1140 (1998).
[23] T. Ogawa, K. Kameyama, R. Kuc, and Y. Kosugi, ''Source localization with network inversion using an answer-in-weights scheme,'' IEICE Trans. Inf. Syst. **E79-D**, 608–619 (1996).
[24] W. W. L. Au, ''Comparison of sonar discrimination—dolphin and artificial neural network,'' J. Acoust. Soc. Am. **95**, 2728–2735 (1994).
[25] W. Chang, B. Bosworth, and G. C. Carter, ''Results of using an artificial neural network to distinguish single echoes from multiple sonar echoes,'' J. Acoust. Soc. Am. **94**, 1404–1408 (1993).
[26] V. Tonard and J. Chatillon, ''Acoustical imaging of extended targets by means of synthetic-aperture sonar technique,'' Acustica **83**, 992–999 (1997).
[27] S. Stergiopoulos, ''Implementation of adaptive and synthetic-aperture processing schemes in integrated active-passive sonar systems,'' Proc. IEEE **86**, 358–396 (1998).
[28] J. Zemanek, ''Beam behavior within the nearfield of a vibrating piston,'' J. Acoust. Soc. Am. **49**, 181–191 (1971).
[29] A. D. Pierce, *Acoustics, An Introduction to Its Physical Principles and Applications* (McGraw-Hill, New York, 1981).
[30] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).
[31] L. W. Camp, *Underwater Acoustics* (Wiley-Interscience, New York, 1970), Chap. 7, p. 166.
[32] Ö. Bozma and R. Kuc, ''Characterizing pulses reflected from rough surfaces using ultrasound,'' J. Acoust. Soc. Am. **89**, 2519–2531 (1991).
[33] B. Ayrulu and B. Barshan, ''Identification of target primitives with multiple decision-making sonars using evidential reasoning,'' Int. J. Robotics Res. **17**, 598–623 (1998).
[34] B. Ayrulu, ''Classification of target primitives with sonar using two nonparametric data-fusion methods,'' Master's thesis, Bilkent University, Department of Electrical Engineering, Ankara, Turkey, July 1996.
[35] B. Barshan and R. Kuc, ''A bat-like sonar system for obstacle localization,'' IEEE Trans. Syst. Man Cybern. **22**, 636–646 (1992).
[36] B. Barshan, *A sonar-based mobile robot for batlike prey capture*, Ph.D. thesis, Yale University, Department of Electrical Engineering, New Haven, CT, December 1991. University of Michigan Microfilms, order number 9224325.
[37] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in Pascal* (Cambridge University Press, Cambridge, 1989), pp. 574–579.
[38] B. Barshan and B. Ayrulu, ''Performance comparison of four methods of time-of-flight estimation for sonar waveforms,'' Electron. Lett. **34**, 1616–1617 (1998).
[39] Polaroid Corporation, ''Ultrasonic components group,'' 119 Windsor St., Cambridge, MA 02139 (1990).
[40] B. Barshan and O. Arıkan, ''Performance analysis of two linear array processing algorithms for point-obstacle localization,'' in *Proceedings SPIE Signal and Data Processing of Small Targets, San Diego, CA, July 11–13, 1995*, edited by O. E. Drummond (SPIE, Bellingham, WA, 1995), Vol. 2561, pp. 533–544.
[41] M. L. Hong and L. Kleeman, ''Analysis of ultrasonic differentiation of three-dimensional corners, edges and planes,'' in *Proceedings IEEE International Conference on Robotics and Automation, Nice, France, May 12–*

14, 1992 (IEEE Computer Society Press, Los Alamitos, CA 1992), pp. 580–584.
[42] J. J. Leonard and H. F. Durrant-Whyte, ''Mobile robot localization by tracking geometric beacons,'' IEEE Trans. Rob. Autom. **7**, 376–382 (1991).
[43] B. Barshan and R. Kuc, ''Differentiating sonar reflections from corners and planes by employing an intelligent sensor,'' IEEE Trans. Pattern. Anal. Mach. Intell. **12**, 560–569 (1990).
[44] R. Kuc and M. W. Siegel, ''Physically-based simulation model for acoustic sensor robot navigation,'' IEEE Trans. Pattern. Anal. Mach. Intell. **PAMI-9**, 766–778 (1987).
[45] D. Başkent and B. Barshan, ''Morphological surface profile extraction from multiple sonars,'' in *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems, Victoria, B.C.*, Canada, *October 1998* (IEEE, Piscataway, NJ, 1998), pp.1515–1520.
[46] D. Başkent and B. Barshan, ''Surface profile determination from multiple sonar data using morphological processing,'' Int. J. Robotics Res., in press.

Separation of an acoustic signal from noise via the use of point pressure and velocity measurements

Peter R. Stepanishen
Department of Ocean Engineering, University of Rhode Island, Narragansett, Rhode Island 02882-1197

(Received 3 September 1998; accepted for publication 15 January 1999)

The separation of an acoustic plane wave signal from noise via the use of pressure and velocity measurements at a point is investigated using a time domain and Fourier transform approach. The noise is assumed to originate from a compact source at a known location. A time domain analysis of the separation method for the deterministic signal problem is first presented. A spectral analysis which provides additional insight into the method is then presented. An analysis of the sensitivity of the signal separation method to processor errors including quantization errors and extraneous uncorrelated sensor noise is also presented. Numerical results for some simple examples illustrate the use of a digital signal processor for separating signal from narrow-band and broadband noise. The sensitivity of the signal separation method to processor errors including numerical differentiation and quantization errors is simply illustrated via the examples. In addition, the sensitivity of the separation method to the noise source/sensor separation distance is also presented.
© *1999 Acoustical Society of America.* [S0001-4966(99)04304-0]

PACS numbers: 43.58.Fm, 43.38.Ar [SLE]

INTRODUCTION

The separation of a signal from noise is a classical problem of interest within the sonar community. For the case where the signal source and noise sources are spatially separated, directive sensors and/or arrays of sensors may be used to perform spatial filtering[1,2] to separate the signal and noise for either nonoverlapping or overlapping frequency spectrums. When the signal source and noise sources are collinear, spatial filtering or beamforming is of less value; however, the signal and noise can again be separated via the use of frequency filtering if the signal and noise spectrums are nonoverlapping.[2,3] Finally, when the signal source and noise sources are collinear and their spectrums are overlapping the signal separation problem is more problematic.

Consider now the separation of an acoustic plane wave signal from the noise field generated by a compact noise source, as illustrated in Fig. 1. As a result of recent interest in the development of acoustic particle velocity sensors,[4,5] a method is presented to separate an acoustic plane wave signal from noise via the use of pressure and velocity measurements at a single point. The method was previously proposed by A. Parvulescu[6] in 1961 for the collinear case where $\theta = 0$. Some experiments which confirmed the method were reported shortly thereafter by Lord.[7]

An analysis of the proposed signal and noise separation method for broadband signal and noise sources is presented here using a time domain and a Fourier transform approach. The spectral analysis provides considerable insight into the limitations of the general separation method. Extraneous additive noise associated with the pressure and velocity measurements at a single point is also introduced to provide a measure of the sensitivity of the separation method to such noise. Relatively simple expressions for the errors in the separation method are developed in both the time and frequency domains.

Numerical results are then presented to illustrate the important limitations on the use of the method for separating signal from noise via the use of a simple digital signal processor. Both narrow-band and broadband noises sources are addressed. In particular, the sensitivity of the signal separation method to the following factors is presented: sampling interval, discrete approximation for a numerical differentiator in the digital signal processor, separation distance between the noise source and the pressure/velocity sensor, and additive extraneous uncorrelated pressure and velocity noise.

I. THEORY

Consider a fluid with a compact noise source and an acoustic plane wave signal as illustrated in Fig. 1. The general problem of interest is to separate the signal from the noise for the case where the location of the noise source is not collinear with the direction of the incoming plane wave. For the sake of simplicity a two-dimensional problem is addressed here where the noise source is assumed to be a monopole source. The coordinate system of interest is determined by the known location of the noise source and the sensor which consists of an idealized point hydrophone and a biaxial acoustic velocity sensor which can independently measure particle velocity components in the x_1 and x_2 directions.

In a source free fluid the acoustic pressure $p(\mathbf{x},t)$ satisfies the following linearized wave equation:

$$\nabla^2 p(\mathbf{x},t) = \frac{1}{c_0^2}\frac{\partial^2}{\partial t^2} p(\mathbf{x},t), \tag{1}$$

where c_0 is the acoustic wave speed. With little loss in generality, the monopole noise pressure field $p^N(\mathbf{x},t)$ can be represented as follows:

$$p^N(\mathbf{x},t) = \frac{r_0}{r} p^N(t-(r-r_0)/c_0), \tag{2}$$

 0001-4966/99/105(4)/2332/10/$15.00 © 1999 Acoustical Society of America

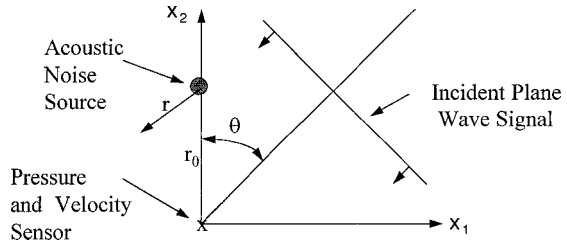


FIG. 1. Geometry for the basic problem of interest.

where r is the distance from the source and r_0 is the known distance to the receiver as noted in Fig. 1. For the case of a plane wave signal propagating at the angle θ relative to the x_2 axis, the signal pressure $p^s(\mathbf{x},t)$ can be expressed as:

$$p^s(\mathbf{x},t)=p^s(t+[x_1\sin\theta+x_2\cos\theta]/c_0). \tag{3}$$

The total pressure field of interest can then be expressed as:

$$p(\mathbf{x},t)=p^s(\mathbf{x},t)+p^N(\mathbf{x},t). \tag{4}$$

Now the associated velocity field $\mathbf{v}(\mathbf{x},t)$ can be obtained from the linearized momentum equation:

$$\rho_0\frac{\partial\mathbf{v}(\mathbf{x},t)}{\partial t}=-\nabla p(\mathbf{x},t), \tag{5}$$

where ρ_0 is the density of the fluid. For the two-dimensional problem of interest, it then follows that:

$$\mathbf{v}(\mathbf{x},t)=\mathbf{v^s}(\mathbf{x},t)+\mathbf{v^N}(\mathbf{x},t), \tag{6}$$

$$=v_1(\mathbf{x},t)\widehat{x_1}+v_2(\mathbf{x},t)\widehat{x_2}. \tag{7}$$

Furthermore, at the sensor location $\mathbf{x}=0$, it is apparent from symmetrical considerations that:

$$\begin{aligned} v_1(0,t)&=v_1^s(0,t),\\ v_2(0,t)&=v_2^s(0,t)+v_2^N(0,t), \end{aligned} \tag{8}$$

where

$$\begin{aligned} v_1^s(0,t)&=-\frac{\sin\theta}{\rho_0c_0}p^s(t),\\ v_2^s(0,t)&=-\frac{\cos\theta}{\rho_0c_0}p^s(t), \end{aligned} \tag{9}$$

and $v_2^N(0,t)$ remains to be determined.

It is readily apparent from the geometry in Fig. 1 that

$$v_2^N(0,t)=-v_r^N(r,t), \tag{10}$$

where $v_r^N(r,t)$ is the radial velocity relative to the noise source. It follows from Eq. (5) that the radial velocity:

$$\begin{aligned} \rho_0\frac{\partial v_r^N(r,t)}{\partial t}&=-\frac{\partial p^N(\mathbf{x}(\mathbf{r}),t)}{\partial r}\\ &=\frac{r_0}{r^2}p^N(t-(r-r_0)/c_0)\\ &\quad+\frac{r_0}{rc_0}p^{N\prime}(t-(r-r_0)/c_0), \end{aligned} \tag{11}$$

where $p^{N\prime}(\tau)\equiv dp^N(\tau)/d\tau$. After integrating Eq. (11) with respect to t, it then follows that

$$\begin{aligned} v_r^N(r,t)&=\frac{p^N(t-(r-r_0)/c_0))}{\rho_0c_0}\frac{r_0}{r}\\ &\quad+\frac{r_0}{\rho_0r^2}\int_{-\infty}^{t-(r-r_0)/c_0}p^N(t')dt'. \end{aligned} \tag{12}$$

Finally, at $\mathbf{x}=0$ it is noted that

$$v_2^N(0,t)=-\frac{p^N(t)}{\rho_0c_0}-\frac{1}{\rho_0r_0}\int_{-\infty}^{t}p^N(t')dt'. \tag{13}$$

As noted earlier, the sensor consists of an idealized hydrophone and a biaxial acoustic velocity sensor. Since the sensor is an idealized point sensor, the hydrophone response is omnidirectional for all frequencies of interest; however, the associated biaxial acoustic velocity sensor, which can independently measure particle velocity components in the x_1 and x_2 directions, is inherently directional for all frequencies. More specifically, it is noted from Eq. (9) that the directional response of the ideal particle velocity sensor oriented along the x_1 axis is $\cos\theta$, whereas the directional response of the ideal particle velocity sensor oriented along the x_2 axis is $\sin\theta$. The pressure and velocity sensitivities are normalized to unity for convenience.

The outputs for the idealized pressure and particle velocity sensor can now be simply expressed in terms of $p^s(t)$ and $p^N(t)$, i.e., the signal and noise pressures at the sensor. More specifically, it now follows from Eqs. (4) and (8) that:

$$\begin{aligned} &p^s(t)+p^N(t)=p(0,t),\\ &-\frac{\sin\theta}{\rho_0c_0}p^s(t)=v_1(0,t),\\ &-\frac{p^N(t)}{\rho_0c_0}-\frac{1}{\rho_0r_0}\int_{-\infty}^{t}p^N(t')dt'-\frac{\cos\theta}{\rho_0c_0}p^s(t)=v_2(0,t). \end{aligned} \tag{14}$$

Since the range to the noise source r_0 is known, there are three unknowns [$p^s(t)$, $p^N(t)$ and θ] and three equations in Eq. (14). Several simple solutions of Eq. (14) for special cases are thus possible, as noted in the following paragraphs.

If the angle of the incoming plane wave signal θ is known, it is readily apparent from Eq. (14) that the signal and noise can in principle be separated as follows:

$$p^s(t)=-\frac{\rho_0c_0}{\sin\theta}v_1(0,t) \tag{15}$$

and

$$p^N(t)=p(0,t)+\frac{\rho_0c_0}{\sin\theta}v_1(0,t). \tag{16}$$

Clearly no information on the location of the noise source is required to determine $p^s(t)$ from Eq. (15) for a known nonzero θ if no extraneous noise is present in $v_1(0,t)$. It is also clearly evident from Eq. (16) that the determination of $p^N(t)$ is affected by extraneous noise in $p(0,t)$. Since θ and the noise location r_0 are known, the remaining equation in Eq. (14) can then be used to provide a consistency check or a measure of the accuracy of the estimates for the two unknowns, i.e., $p^s(t)$ and $p^N(t)$.

Although Eqs. (15) and (16) can be used to determine $p^s(t)$ and $p^N(t)$ for $\theta \neq 0$, there is clearly a problem for the case where $\theta = 0°$. For this case it is apparent from Eq. (14) that the set of three equations reduce to the following two equations:

$$p^s(t) + p^N(t) = p(0,t), \tag{17}$$

$$\frac{-p^s(t)}{\rho_0 c_0} - \frac{1}{\rho_0 c_0}\left[p^N(t) + \frac{c_0}{r_0}\int_{-\infty}^{t} p^N(\tau)\,d\tau\right] = v_2(0,t), \tag{18}$$

where $p(0,t)$ and $v_2(0,t)$ are the known sensor outputs and $p^s(t)$ and $p^N(t)$ are to be determined. The reduction in the number of equations for $\theta = 0$ is of course to be expected based on the directional characteristics of the velocity sensors as noted above.

The separation of signal and noise for the collinear case where $\theta = 0$ is now simply performed. After substituting Eq. (17) into Eq. (18) and rearranging it follows that:

$$\frac{c_0}{r_0}\int_{-\infty}^{t} p^N(\tau)\,d\tau = p(0,t) + \rho_0 c_0 v_2(0,t). \tag{19}$$

After differentiating Eq. (19) with respect to t, it is apparent that

$$p^N(t) = \frac{r_0}{c_0}\frac{d}{dt}[p(0,t) + \rho_0 c_0 v_2(0,t)]. \tag{20}$$

The signal of interest can then be expressed as:

$$\begin{aligned} p^s(t) &= p(0,t) - p^N(t) \\ &= p(0,t) - \frac{r_0}{c_0}\frac{d}{dt}[p(0,t) + \rho_0 c_0 v_2(0,t)]. \end{aligned} \tag{21}$$

Irrespective of their spectral properties, the signal and noise have been separated via the use of the measured quantities $p(0,t)$ and $v_2(0,t)$.

The effects of extraneous noise on the signal separation method are now addressed via the addition of extraneous noise to the measured pressure and velocity components. More specifically, the pressure and particle velocity measurements are now represented as:

$$\begin{aligned} &p(0,t) \rightarrow p(0,t) + p^n(0,t), \\ &v_1(0,t) \rightarrow v_1(0,t) + v_1^n(0,t), \\ &v_2(0,t) \rightarrow v_2(0,t) + v_2^n(0,t). \end{aligned} \tag{22}$$

As a result of the extraneous noises, the signal and noise pressures are now represented as:

$$p^s(t) \rightarrow \widehat{p^s(t)}, \quad p^N(t) \rightarrow \widehat{p^N(t)}. \tag{23}$$

It then follows that Eq. (14) may now be expressed as:

$$\begin{aligned} &\widehat{p^s(t)} + \widehat{p^N(t)} = p(0,t) + p^n(0,t), \\ &-\frac{\sin\theta}{\rho_0 c_0}\widehat{p^s(t)} = v_1(0,t) + v_1^n(0,t), \\ &-\frac{\widehat{p^N(t)}}{\rho_0 c_0} - \frac{1}{\rho_0 r_0}\int_{-\infty}^{t}\widehat{p^N(t')}\,dt' - \frac{\cos\theta}{\rho_0 c_0}\widehat{p^s(t)} \\ &\quad = v_2(0,t) + v_2^n(0,t). \end{aligned} \tag{24}$$

For the case of $\theta \neq 0$ it is apparent from Eq. (24) that:

$$\begin{aligned} &\widehat{p^s(t)} = -\frac{\rho_0 c_0}{\sin\theta}[v_1(0,t) + v_1^n(0,t)], \\ &\widehat{p^N(t)} = p(0,t) + p^n(0,t) + \frac{\rho_0 c_0}{\sin\theta}[v_1(0,t) + v_1^n(0,t)]. \end{aligned} \tag{25}$$

After using Eqs. (15) and (16) it then follows that:

$$\begin{aligned} &\widehat{p^s(t)} = p^s(t) - \frac{\rho_0 c_0}{\sin\theta}v_1^n(0,t), \\ &\widehat{p^N(t)} = p^N(t) + p^n(0,t) + \frac{\rho_0 c_0}{\sin\theta}v_1^n(0,t). \end{aligned} \tag{26}$$

It is clearly apparent that $\widehat{p^s(t)} = p^s(t)$ and $\widehat{p^N(t)} = p^N(t)$ for the noise free case; however, large errors can result in the estimation of $p^s(t)$ for $v_1^n(0,t) \neq 0$ and small θ. This error plus those associated with $p^n(0,t)$ clearly affect the estimation of $p^N(t)$.

The separation of signal and noise for the collinear case with extraneous noise where $\theta = 0$ is also simply performed. In a procedure analogous to that followed for the extraneous noise free case, it is easily shown that:

$$\begin{aligned} &\widehat{p^s(t)} + \widehat{p^N(t)} = p(0,t) + p^n(0,t), \\ &\frac{-\widehat{p^s(t)}}{\rho_0 c_0} - \frac{1}{\rho_0 c_0}\left[\widehat{p^N(t)} + \frac{c_0}{r_0}\int_{-\infty}^{t}\widehat{p^N(\tau)}\,d\tau\right] \\ &\quad = v_2(0,t) + v_2^n(0,t). \end{aligned} \tag{27}$$

It then follows that the estimated signal and noise can be expressed as:

$$\begin{aligned} \widehat{p^s(t)} &= p(0,t) + p^n(0,t) + \frac{r_0}{c_0}\frac{d}{dt}[p(0,t) + p^n(0,t) \\ &\qquad + \rho_0 c_0 v_2(0,t) + \rho_0 c_0 v_2^n(0,t)], \\ \widehat{p^N(t)} &= p(0,t) + p^n(0,t) - \widehat{p^s(t)}. \end{aligned} \tag{28}$$

A block diagram of a signal processor to perform the indicated operations in Eqs. (28) is shown in Fig. 2.

The effect of the extraneous noise on the signal separation method for $\theta = 0$ can now be simply obtained for the case of an ideal differentiator via the substitution of Eqs. (19) and (20) into (28), which leads to:

$$\begin{aligned} &\widehat{p^s(t)} = p^s(t) + p^n(0,t) + \frac{r_0}{c_0}\frac{d}{dt}[p^n(0,t) + \rho_0 c_0 v_2^n(0,t)], \\ &\widehat{p^N(t)} = p(0,t) + p^n(0,t) - \widehat{p^s(t)}. \end{aligned} \tag{29}$$

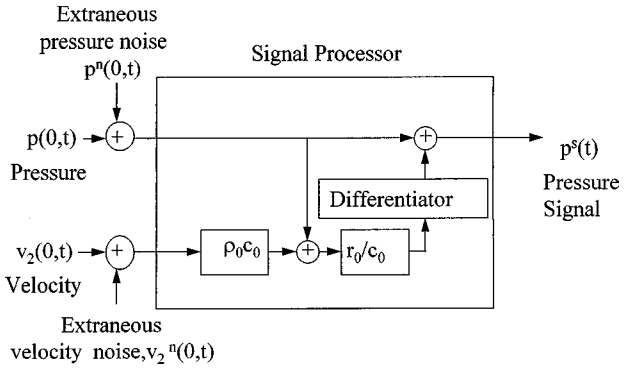


FIG. 2. Signal processor to separate the signal from a noise.

It then follows that the associated errors in estimating the signal and noise for the case of an ideal differentiator can be expressed as:

$$\begin{aligned}\epsilon^s(t) &= p^n(0,t) + \frac{r_0}{c_0}\frac{d}{dt}[p^n(0,t) + \rho_0 c_0 v_2^n(0,t)],\\ \epsilon^N(t) &= -\frac{r_0}{c_0}\frac{d}{dt}[p^n(0,t) + \rho_0 c_0 v_2^n(0,t)]\\ &= -\epsilon^s(t) + p^n(0,t).\end{aligned} \tag{30}$$

If an ideal differentiator is not used, e.g., due to a digital implementation of the processor in Fig. 2, additional errors will occur in estimating the signal and noise as noted in the following section.

In order to obtain further insight into the errors and limitations of the signal separation method, spectral methods are now employed. After introducing the following Fourier transform pair, i.e., $p(\mathbf{x},t) \Leftrightarrow P(\mathbf{x},\Omega)$ where:

$$P(\mathbf{x},\Omega) = \int_{-\infty}^{\infty} p(\mathbf{x},t) e^{-i\Omega t}\, dt, \tag{31}$$

$$p(\mathbf{x},t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} P(\mathbf{x},\Omega) e^{i\Omega\tau}\, d\Omega, \tag{32}$$

the following transformed wave equation is then simply obtained:

$$[\nabla^2 + K^2] P(\mathbf{x},\Omega) = 0, \tag{33}$$

where $K = \Omega/c_0$.

In order to investigate the separation of signal and noise in the spectral domain, the following set of spectral equations is readily obtained from the Fourier transform of Eq. (24):

$$\begin{aligned}&\widehat{P^s(\Omega)} + \widehat{P^N(\Omega)} = P(0,\Omega) + P^n(0,\Omega),\\ &\frac{-\sin\theta}{\rho_0 c_0}\widehat{P^s(\Omega)} = V_1(0,\Omega) + V_1^n(0,\Omega),\\ &-\frac{\cos\theta}{\rho_0 c_0}\widehat{P^s(\Omega)} - \frac{\widehat{P^N(\Omega)}}{Z_a(Kr_0)} = V_2(0,\Omega) + V_2^n(0,\Omega),\end{aligned} \tag{34}$$

where the specific acoustic impedance $Z_a(Kr_0)$ is

$$Z_a(Kr_0) = \rho_0 c_0 \frac{jKr_0}{1 + jKr_0}. \tag{35}$$

For $Kr_0 \gg 1$ it is easily seen that

$$Z_a(Kr_0) \sim \rho_0 c_0. \tag{36}$$

The separation of signal and noise in the spectral domain for the case of $\theta \neq 0$ may be simply obtained via the Fourier transform of Eq. (24) or via Eq. (34), which lead to:

$$\begin{aligned}\widehat{P^s(\Omega)} &= P^s(\Omega) - \frac{\rho_0 c_0}{\sin\theta} V_1^n(0,\Omega),\\ \widehat{P^N(\Omega)} &= P^n(\Omega) + P^n(0,\Omega) + \frac{\rho_0 c_0}{\sin\theta} V_1^n(0,\Omega).\end{aligned} \tag{37}$$

It is again no surprise that large spectral errors can result in the estimation of $P^s(\Omega)$ for small θ and nonzero $V_1^n(0,\Omega)$. Once again, these errors plus those associated with $P^n(0,\Omega)$ clearly affect the estimation of $P^n(\Omega)$.

The spectral separation of a signal from noise when $\theta = 0$ is also simply addressed via the use of Eq. (34) which lead to the following set of linear algebraic equations:

$$\begin{bmatrix} 1 & 1 \\ \dfrac{-1}{\rho_0 c_0} & \dfrac{-1}{Z_a(Kr_0)} \end{bmatrix} \begin{bmatrix} \widehat{P^s(\Omega)} \\ \widehat{P^N(\Omega)} \end{bmatrix} = \begin{bmatrix} P(0,\Omega) + P^n(0,\Omega) \\ V_2(0,\Omega) + V_2^n(0,\Omega) \end{bmatrix}. \tag{38}$$

It then follows by matrix inversion that

$$\begin{bmatrix} \widehat{P^s(\Omega)} \\ \widehat{P^N(\Omega)} \end{bmatrix} = \frac{1}{\Delta}\begin{bmatrix} \dfrac{-1}{Z_a(Kr_0)} & -1 \\ \dfrac{1}{\rho_0 c_0} & 1 \end{bmatrix} \begin{bmatrix} P(0,\Omega) + P^n(0,\Omega) \\ V_2(0,\Omega) + V_2^n(0,\Omega) \end{bmatrix}, \tag{39}$$

where the determinant Δ is

$$\Delta = \frac{-1}{jKr_0\rho_0 c_0}. \tag{40}$$

As $Kr_0 \to \infty$ it is obvious that $\Delta \to 0$. Although the linear algebraic equations in Eq. (38) are then linearly related for the extraneous noise free case, the equations are generally inconsistent for the noisy case. It is thus apparent that problems can exist with the signal separation method for spectral components where $Kr_0 \gg 1$.

After some algebra based on Eq. (39), the estimated signal $\widehat{P^s(\Omega)}$ and noise $\widehat{P^N(\Omega)}$ can be simply expressed as:

$$\begin{aligned}\widehat{P^s(\Omega)} &= P^s(\Omega) + P^n(0,\Omega) + \frac{j\Omega r_0}{c_0}[P^n(0,\Omega)\\ &\quad + \rho_0 c_0 V_2^n(0,\Omega)],\\ \widehat{P^N(\Omega)} &= P^N(\Omega) - \frac{j\Omega r_0}{c_0}[P^n(0,\Omega) + \rho_0 c_0 V_2^n(0,\Omega)].\end{aligned} \tag{41}$$

These results are of course simply equivalent to the Fourier transform of Eq. (29). The associated errors in estimating the signal and noise for this case, which are implicitly based on the use of an ideal differentiator, can then be expressed as:

$$\epsilon^s(\Omega)=P^n(0,\Omega)+\frac{j\Omega r_0}{c_0}[P^n(0,\Omega)+\rho_0 c_0 V_2^n(0,\Omega)],$$
$$\epsilon^N(\Omega)=-\frac{j\Omega r_0}{c_0}[P^n(0,\Omega)+\rho_0 c_0 V_2^n(0,\Omega)] \tag{42}$$
$$=-\epsilon^s(\Omega)+p^n(0,\Omega),$$

and they illustrate that the high frequency extraneous noise components where $Kr_0 \gg 1$ can be expected to result in problems with the signal separation method.

II. SOME EXAMPLES OF INTEREST

Numerical results are now presented to illustrate the effectiveness of the method presented in the previous section to separate a signal from a collinear noise source where $\theta=0$ in Fig. 1. The signal processor to separate the signal and noise is based on the model shown in Fig. 2. In contrast to the previous section, both ideal and nonideal differentiators[8,9] are now considered where the latter are based on a simple first order finite difference approximation. Although improved higher order finite difference approximations are available, the present results indicate the sensitivity of the noise separation method to the extraneous noise and the errors associated with a simple finite difference approximation for the differentiator.

The extraneous noise sources, $p^n(0,t)$ and $v_2^n(0,t)$, are considered here to be sample functions of zero mean Gaussian ergodic random processes. For the purpose of normalizing the numerical results, an equivalent noise pressure for the velocity is introduced as follows:

$$p_v^n(0,t)=\rho_0 c_0 v_2^n(0,t). \tag{43}$$

For simplicity, both random processes $\{p^n(0,t)\}$ and $\{p_v^n(0,t)\}$ are assumed to have either uniform or Gaussian probability distributions with zero means and identical variances denoted by σ^2.

In light of the above definitions it is apparent from Eq. (28) that $\widehat{p^s(t)}$ and $\widehat{p^N(t)}$ can be determined from:

$$\widehat{p^s(t)}=p(0,t)+p^n(0,t)+\frac{r_0}{c_0}\frac{d}{dt}[p(0,t)+p^n(0,t)$$
$$+\rho_0 c_0 v_2(0,t)+p_v^n(0,t)],$$
$$\widehat{p^N(t)}=p(0,t)+p^n(0,t)-\widehat{p^s(t)}. \tag{44}$$

For the case of an ideal differentiator, the associated errors in estimating the signal and noise can then be expressed as:

$$\epsilon^s(t)=p^n(0,t)+\frac{r_0}{c_0}\frac{d}{dt}[p^n(0,t)+p_v^n(0,t)],$$
$$\epsilon^N(t)=-\frac{r_0}{c_0}\frac{d}{dt}[p^n(0,t)+p_v^n(0,t)]. \tag{45}$$

It is thus apparent that the errors in the separation method, which are associated with the extraneous noise sources $p^n(0,t)$ and $p_v^n(0,t)$, will generally increase as r_0 increases and the bandwidth of the noise sources increases.

In addition to the errors associated with the extraneous noise sources, it is apparent that errors can arise via the use of any approximation for the differentiator in Fig. 2. In order to illustrate the nature of these errors for a simple digital processor, the following approximation based on a simple first order finite difference approximation is noted:

$$\frac{dg(t)}{dt}\approx\frac{1}{\Delta t}[g(t)-g(t-\Delta t)]\equiv g_1(t). \tag{46}$$

In contrast to the ideal differentiator in which

$$\frac{dg(t)}{dt}\Leftrightarrow j\Omega G(\Omega), \tag{47}$$

it is easily shown that:

$$g_1(t)\Leftrightarrow j\Omega G(\Omega)e^{-j\Omega\Delta t/2}\frac{\sin\Omega\Delta t/2}{\Omega\Delta t/2}, \tag{48}$$
$$\approx j\Omega G(\Omega)\{1-j\Omega\Delta t/2\},\quad \Omega\Delta t\ll 1. \tag{49}$$

It is then apparent for band-limited signals where Ω_M is the maximum frequency and $\Omega_M\Delta t\ll 1$ that

$$g_1(t)\approx\frac{dg(t)}{dt}-\frac{\Delta t}{2}\frac{d^2g(t)}{dt^2}. \tag{50}$$

The errors associated with the use of the simple first order finite difference approximation for the differentiator in Fig. 2 can now be obtained by replacing the derivative operator in Eq. (44) by $g_1(t)$. for the extraneous noise free case it then follows that:

$$\epsilon^s(t)=-\frac{\Delta t}{2}\frac{dp^N(0,t)}{dt},\quad \epsilon^N(t)=-\epsilon^s(t). \tag{51}$$

The additional errors which arise from the use of the nonideal differentiator can thus be controlled via the sampling interval Δt. In contrast to the results in Eq. (45), the errors associated with the approximation for the differentiator are independent of r_0, the position of the noise source. It is also noted that the errors are independent of $p^s(t)$.

As a result of the general interest in the estimated signal, the remaining focus is now on $\widehat{p^s(t)}$ and the errors in estimating $p^s(t)$ resulting from the use of nonideal differentia-

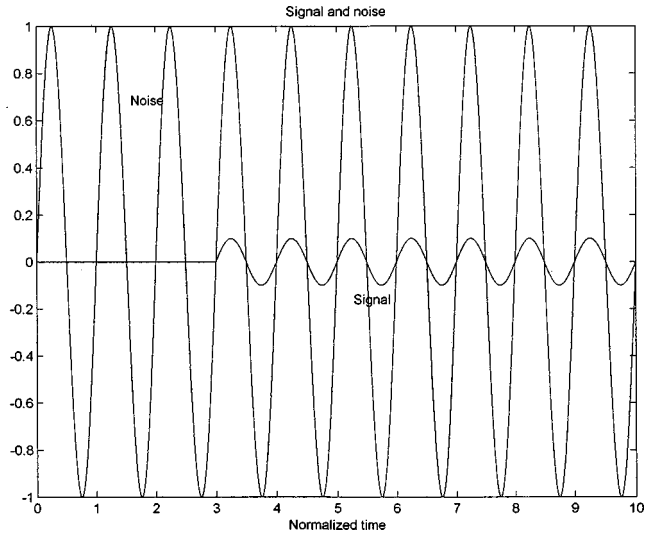


FIG. 3. The signal and noise of interest—−20 dB signal-to-noise ratio.

tors and extraneous noise. After discretizing time and using the first order difference approximation for the derivative, it then follows from Eq. (44) that

$$\begin{aligned}\widehat{p^s(m\Delta t)} &= p(0,m\Delta t)+p^n(0,m\Delta t)+\frac{r_0}{c_0}\frac{1}{\Delta t}[p(0,m\Delta t)\\ &\quad +p^n(0,m\Delta t)+\rho_0 c_0 v_2(0,m\Delta t)\\ &\quad +p_v^n(0,m\Delta t)-[p(0,(m-1)\Delta t)\\ &\quad +p^n(0,(m-1)\Delta t)+\rho_0 c_0 v_2(0,(m-1)\Delta t)\\ &\quad +p_v^n(0,(m-1)\Delta t)]],\\ \widehat{p^N(m\Delta t)} &= p(0,m\Delta t)+p^n(0,m\Delta t)-\widehat{p^s(m\Delta t)}.\end{aligned} \tag{52}$$

For the case of no extraneous noise Eq. (52) then reduces to the following:

$$\begin{aligned}\widehat{p^s(m\Delta t)} &= p(0,m\Delta t)+\frac{r_0}{c_0}\frac{1}{\Delta t}[p(0,m\Delta t)\\ &\quad +\rho_0 c_0 v_2(0,m\Delta t)-[p(0,(m-1)\Delta t)\\ &\quad +\rho_0 c_0 v_2(0,(m-1)\Delta t)]].\end{aligned} \tag{53}$$

Numerical results are now presented for some specific examples of interest. For simplicity, the plane wave signal is defined by the following $p^s(t)$:

$$\begin{aligned}p^s(t) &= P_s \sin[(2\pi f^s(t-T_m)], \quad t>T_m\\ &= 0, \quad t<T_m,\end{aligned} \tag{54}$$

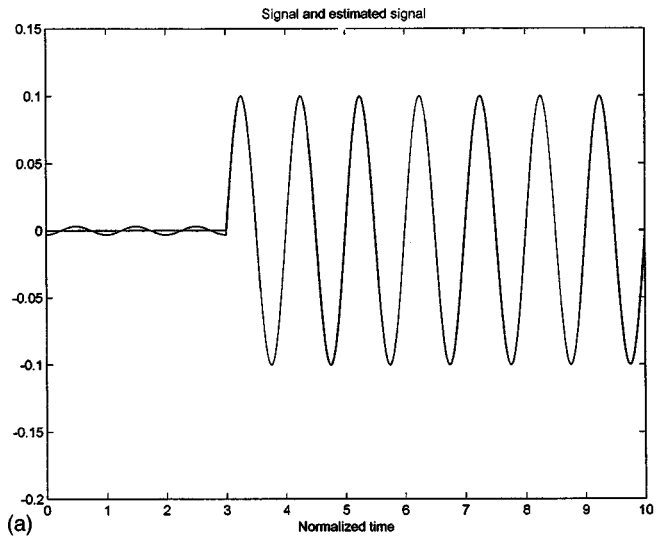
where $T_m = m/f^s$ is a time delay of m periods of the signal frequency f^s. The noise signal of interest $p^N(t)$ is represented here as:

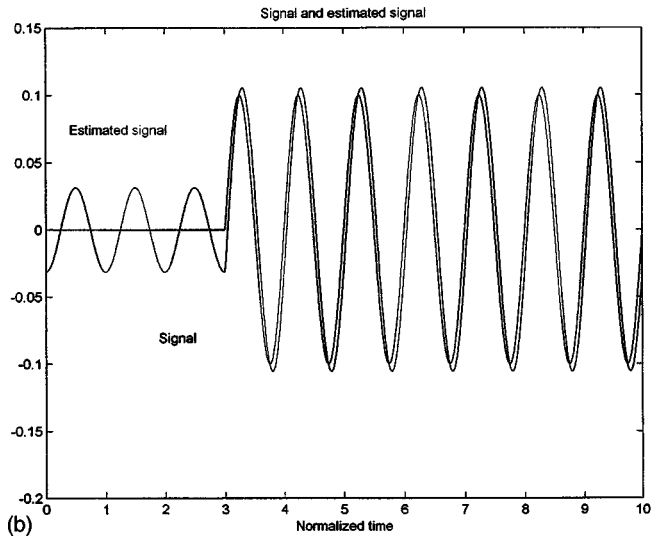
$$p^N(t) = P_N \sin(2\pi f^N t), \quad \text{all } t. \tag{55}$$

It is noted that the noise is assumed to be present for all time whereas the signal is zero up to the time T_m. The signal to noise ratio S/N is defined to be $(P_s/P_N)^2$.

In light of the above definitions it is apparent from Eq. (14) that the signal plus noise at the pressure sensor can be expressed as

$$\begin{aligned}p(0,t) &= P_s \sin[(2\pi f^s(t-T_m)]u(t-T_m)\\ &\quad +P_N \sin(2\pi f^N t),\end{aligned} \tag{56}$$





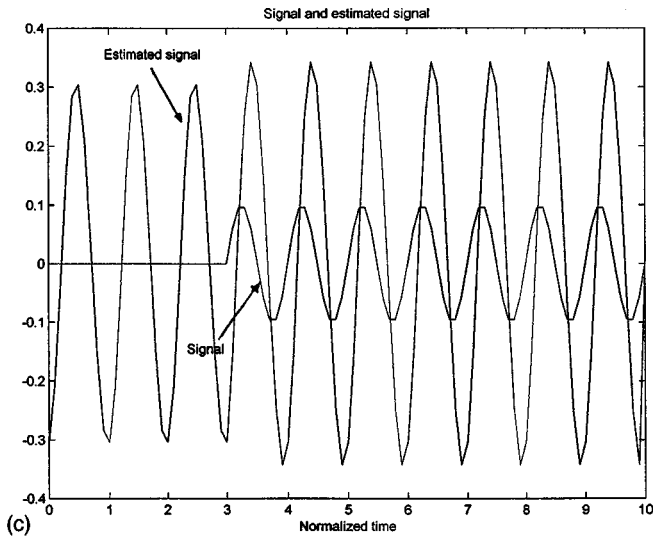


FIG. 4. The signal and estimated signal for $K^N r_0 = 2\pi$. (a) The signal and estimated signal for $K^N r_0 = 2\pi$ and $\Delta t = 0.001$. (b) The signal and estimated signal for $K^N r_0 = 2\pi$ and $\Delta t = 0.01$. (c) The signal and estimated signal for $K^N r_0 = 2\pi$ and $\Delta t = 0.1$.

where $u(\,)$ is the familiar Heaviside function. Similarly, the signal plus noise velocity at the velocity sensor can be simply expressed as follows:

$$v_2(0,t)=\frac{-1}{\rho_0 c_0}\Bigg[P_N\sin(2\pi f^N t) - \frac{c_0}{2\pi f^N r_0}P_N\cos(2\pi f^N t) + P_s\sin[2\pi f^s(t-T_m)]u(t-T_m)\Bigg]. \quad (57)$$

With no loss in generality $P_N=1$ and $f^N=1$ for the following examples with P^s, f^s, and r_0/c_0 variable.

Consider first the case where the extraneous noise is zero, i.e., $p^n(0,t)=p_v^n(0,t)=0$. The noise and the signal of interest are shown in Fig. 3 for the case of $m=3$ with $f^s=f^N=1$ and $P^s=0.1$ which corresponds to a -20 dB signal to noise ratio. The estimated signal is shown in Fig. 4(a), (b), and (c) for $(\Omega^N r_0)/c_0=K^N r_0=2\pi$ and $\Delta t=0.001$, 0.01, and 0.1, respectively. It is clear that the signal is readily apparent in the estimated signal for the case of $\Delta t=0.001$. Although the signal is still clearly observed in the estimated signal for $\Delta t=0.01$, the desired signal is lost for the case of $\Delta t=0.1$. For this latter case the error is significantly greater than the signal of interest.

It is also apparent from the numerical results in Fig. 4 that the error in estimating the signal for the three cases is clearly predicted by Eq. (51) which can be simply expressed as $\epsilon^s(t)=-\pi\Delta t\cos(2\pi t)$. Thus for a specified signal amplitude it is easy to control the relative importance of the error via the selection of Δt. It is also noted from Eq. (51) that the error is independent of the signal and for the present cases the error is in phase quadrature with the signal.

As a second example, consider now the signal and noise for the case of $m=3$ with with $f^s=f^N=1$ and $P^s=1.0$ which corresponds to a 0 dB signal to noise ratio. The sequences of extraneous noise samples $\{p^n(0,m\Delta t)\}$ and $\{p_v^n(0,m\Delta t)\}$ are considered to be nonzero sample sequences of stationary random processes which are assumed to have uniform probability density distributions in the range $[-q/2,q/2]$. It is noted that the uniform probability density distribution is often used in describing the quantization error associated with sampling analog signals.[10] For such cases the quantization error is treated as an additive noise where q corresponds to the quantization width. The sequences of extraneous noise samples or quantization errors are also assumed here to be uncorrelated with the corresponding analog measured sensor outputs $p(0,t)$ and $v_2(0,t)$.

The estimated signal and the signal of interest for $q=0.002$, $K^N r_0=2\pi\times10^{-3}$, and $\Delta t=0.001$ are shown in Fig. 5 and the associated error is shown in Fig. 6. It is apparent that the signal of interest is clearly separated from the noise. It is also apparent that the signal error in Fig. 6 can be viewed as a sum of a deterministic component, which is dependent on Δt as noted above, and a random component associated with the extraneous quantization noise. The estimated signal and the signal of interest for $q=0.002$ and $\Delta t=0.001$ are shown in Figs. 7 and 8 for $K^N r_0=2\pi\times10^{-1}$ and $K^N r_0=2\pi\times10^{+1}$, respectively. These results clearly demonstrate the effect of the sensor position relative to the noise source on the effectiveness of the signal separation method. Although the signal is readily apparent in Fig. 7, the increasing importance of the noise as $K^N r_0$ is increased by a factor of 10 is also visible relative to the results in Fig. 5. Finally, it is evident from the results in Fig. 8 that the signal for this example cannot be separated from the noise for $K^N r_0 \gg 1$. In fact, the results for the estimated signal in Fig. 8, which show little resemblance to the signal, are to be expected in light of the analytical results.

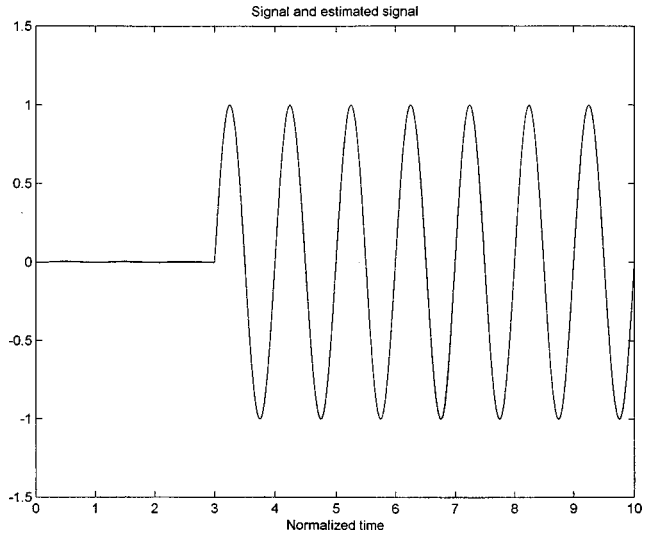


FIG. 5. The signal and the estimated signal for $q=0.002$, $K^N r_0=2\pi\times10^{-3}$, and $\Delta t=0.001$.

For the final class of examples, the same signal is again considered where $m=3$ with $f^s=1$ and $P^s=1.0$. The noise is now specified to be band-limited white Gaussian noise. More specifically, the noise of interest is considered to be a sample function of a zero mean white Gaussian noise process with a variance σ^2 which is passed through a kth order elliptic bandpass filter[10,11] with a passband from f_l to f_h, a passband ripple of R_p dB and a stopband attenuation of R_s

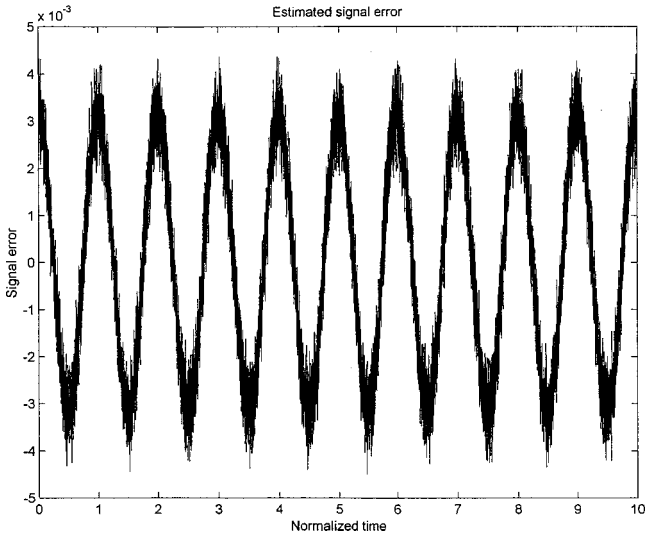


FIG. 6. The error for $q=0.002$, $K^N r_0=2\pi\times10^{-3}$, and $\Delta t=0.001$.

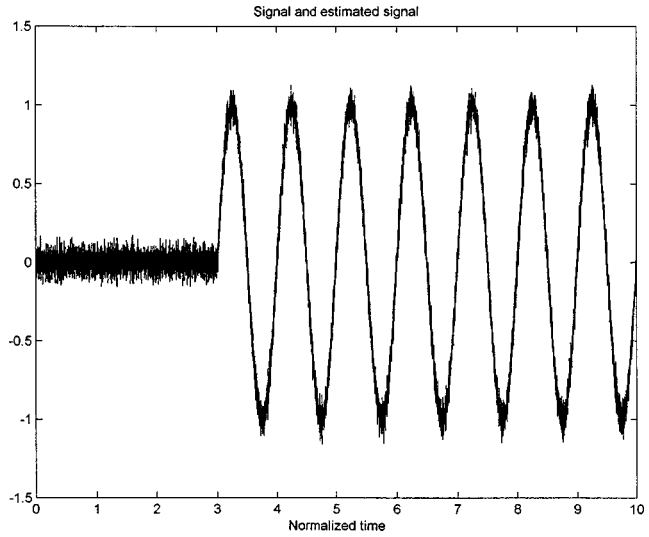


FIG. 7. The signal and the estimated signal for $q=0.002$, $\Delta t=0.001$, and $K^N r_0=2\pi\times10^{-1}$.

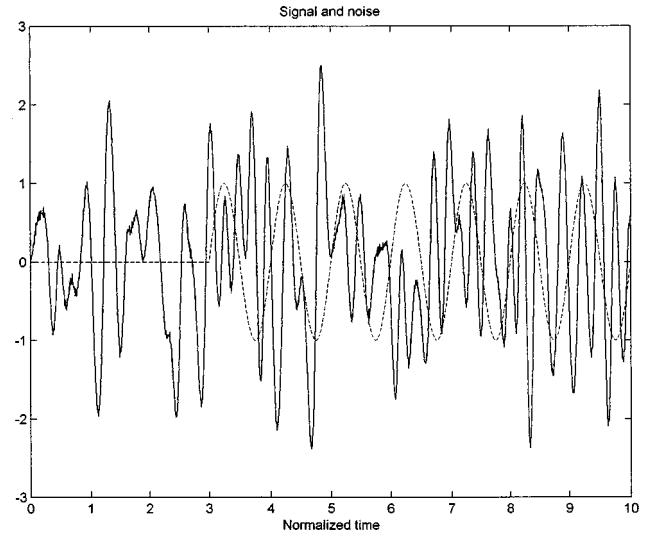


FIG. 9. The signal and zero mean Gaussian bandlimited noise versus time for $\sigma^2=16$, $f_l=0.25$, $f_h=4.0$, and $\Delta t=0.002$.

dB. For the present study, $k=8$, $R_p=0.1$ dB, and $R_s=40$ dB. After discretizing and sampling, the variance of the resulting discrete time random noise process is noted to be approximately $\sigma_n^2=2\sigma^2(f_h-f_l)\Delta t$. For the case where the switch-on time of the sinusoidal signal is ignored, i.e., the signal is assumed to exist for all time, the variance of the signal is $\sigma_s^2=1/2$ and the signal to noise ratio is thus approximately $\sigma_s^2/\sigma_n^2=[4\sigma^2(f_h-f_l)\Delta t]^{-1}$.

Consider first the case where $\sigma^2=16$, $f_l=0.25$, $f_h=4.0$, and $\Delta t=0.002$. The signal and band-limited noise versus time are shown in Fig. 9 where the signal-to-noise ratio is approximately $\sigma_s^2/\sigma_n^2\approx0.83$. The corresponding pressure and velocity signals $p(0,t)$ and $v_2(0,t)$ are presented in Fig. 10 and the estimated signal is presented in Fig. 11 for $r_0/c_0=1000$ with no extraneous noise, i.e., $p^n(0,t)=p_v^n(0,t)=0$. It is clear that the signal is readily apparent in the estimated signal in Fig. 11. If the cutoff frequencies of the filter are now decreased so that $f_l=0.5$ and $f_h=2.0$, the bandwidth is reduced and it is clear from the above discussion that the variance of the noise will also be reduced. The signal and band-limited noise versus time for this case are shown in Fig. 12, the corresponding pressure and velocity signals $p(0,t)$ and $v_2(0,t)$ are presented in Fig. 13 and the estimated signal is presented in Fig. 14. A comparison of the estimated signals in Figs. 11 and 14 clearly shows the expected results. In general, the bandwidth of the filter should of course be matched to that of the signal if it is known *a priori*.

III. SUMMARY AND CONCLUSIONS

A previously proposed method[6] to separate an acoustic signal from noise via the use of pressure and particle velocity measurements at a point in a fluid was investigated theoretically. Both the pressure and bi-axial velocity sensors are assumed to be ideal with matched sensitivities over the fre-

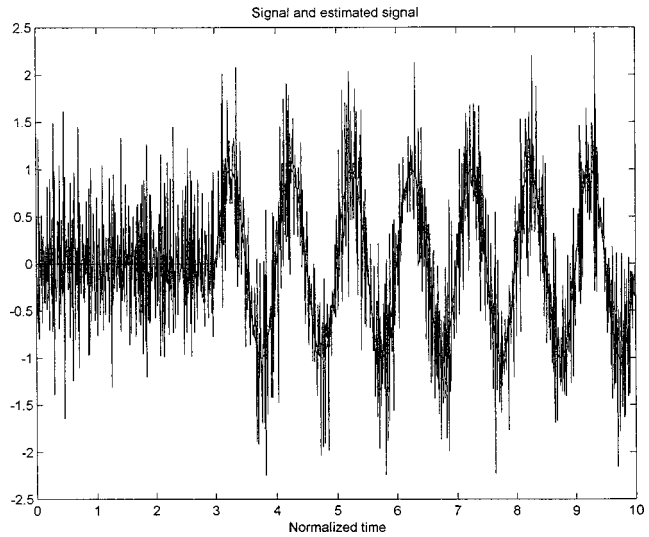


FIG. 8. The signal and the estimated signal for $q=0.002$, $\Delta t=0.001$, and $K^N r_0=2\pi\times10^{+1}$.

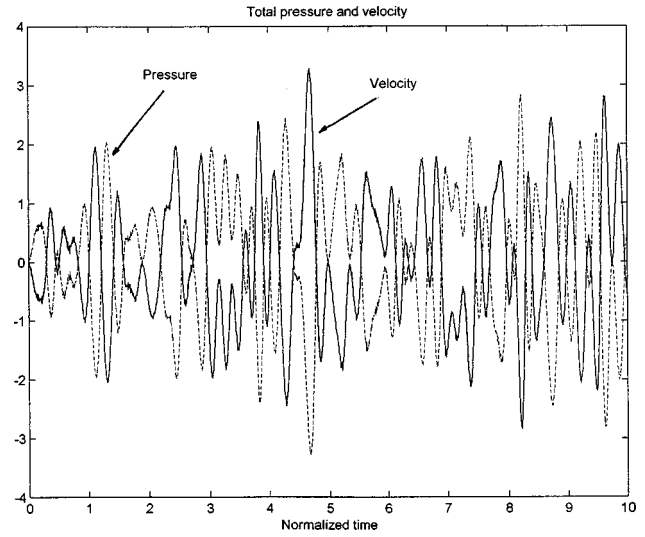


FIG. 10. The pressure and velocity signals $p(0,t)$ and $v_2(0,t)$ for the case of zero mean Gaussian band-limited noise versus time for $\sigma^2=16$, $f_l=0.25$, $f_h=4.0$, and $\Delta t=0.002$.

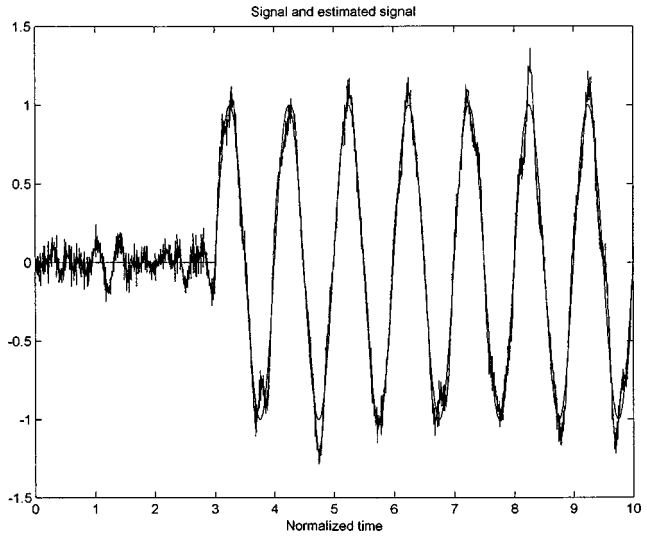


FIG. 11. The estimated signal versus time for $\sigma^2=16$, $f_l=0.25$, $f_h=4.0$, and $\Delta t=0.002$.

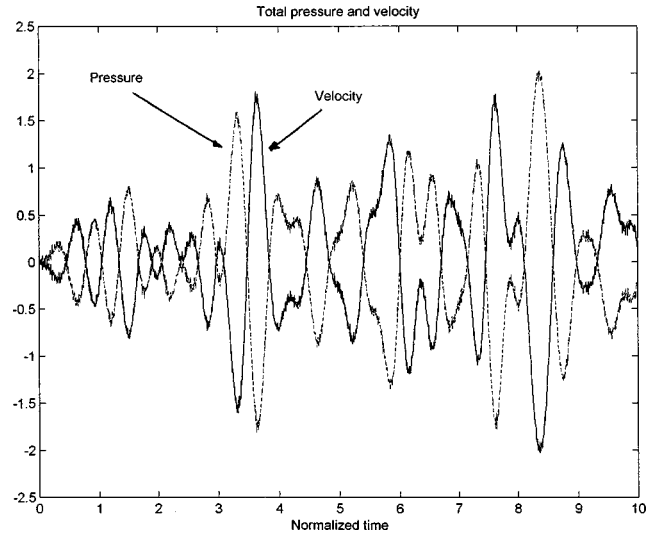


FIG. 13. The pressure and velocity signals $p(0,t)$ and $v_2(0,t)$ for the case of zero mean Gaussian band-limited noise versus time for $\sigma^2=16$, $f_l=0.5$, $f_h=2.0$, and $\Delta t=0.002$.

quency range of interest. In contrast to the pressure sensor which is assumed to be omnidirectional, the bi-axial velocity sensor is assumed to exhibit an ideal dipole directional response.

An analysis of the proposed signal (and noise) separation method for broadband signal and noise sources was presented using time domain and Fourier transform methods. The analysis provides considerable insight into the limitations of the signal separation method using point measurements. In general, a plane wave signal cannot be separated from a noise via the point measurements since there are more unknowns than equations. If the location of the noise source is known and the direction θ of the incoming plane wave signal relative to that of the noise source is known, an overdetermined set of linear equations is obtained. The desired signal can then in principle be simply separated from the noise for $\theta\neq 0°$ via the use of Eqs. (26); however, large errors can be expected in the separation procedure as $\theta\rightarrow 0$.

If the location of the noise source is known and the direction of the incoming plane wave signal is $\theta=0$ a plane wave signal may be separated from the noise via an alternative procedure based on the use of a point pressure and velocity measurement. A simple signal processor to accomplish the separation was presented in Fig. 2. However, it was noted that the separation procedure is based on an ill posed set of algebraic equations for those spectral components of the noise where $\Omega r_0/c_0 \gg 1$. As $\Omega r_0/c_0$ increases, the determinant of the set of algebraic equations approaches zero and the set of equations becomes increasingly poorly conditioned which leads to an increased sensitivity to numerical and/or other noise sources in the separation process. Simply stated, the near field components of the noise source are necessary in order to perform the separation of the signal from the noise.

Numerical results were presented to illustrate the method for separating signal from noise via the use of a simple digital signal processor. The separation of a sinu-

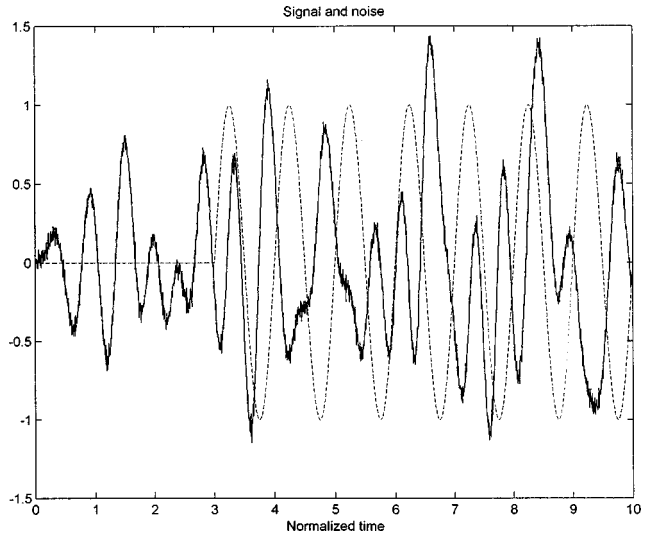


FIG. 12. The signal and zero mean Gaussian band-limited noise versus time for $\sigma^2=16$, $f_l=0.5$, $f_h=2.0$, and $\Delta t=0.002$.

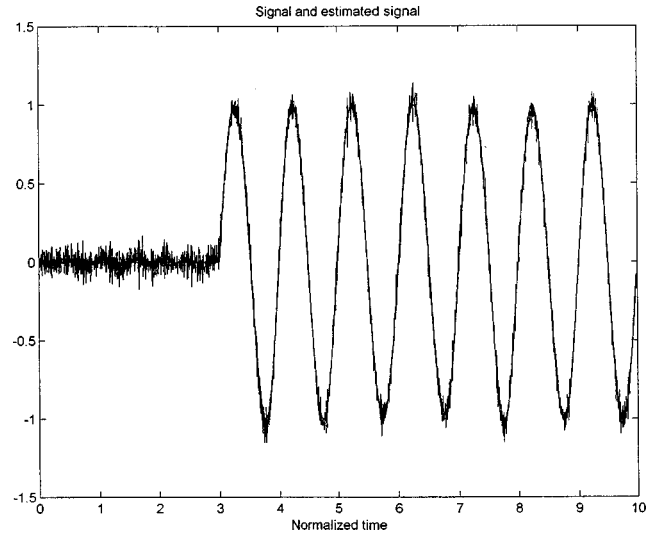


FIG. 14. The estimated signal versus time for $\sigma^2=16$, $f_l=0.5$, $f_h=2.0$, and $\Delta t=0.002$.

soidal signal and noise was first addressed. Although the frequency spectrums of the signal and noise overlap, the signal was readily separated from the noise for a -20 dB signal-to-noise ratio, providing the sampling frequency was large enough and extraneous noise was not present.

The sensitivity of the signal separation method to the separation distance between the noise source and the pressure/velocity sensor, and additive uncorrelated pressure and velocity noise was then addressed. Additive uncorrelated quantization noise was considered as extraneous noise in the digital signal processor. The numerical results clearly demonstrated the important effect of the sensor position relative to the noise source on the effectiveness of the signal separation method. In particular, it was demonstrated that the signal cannot be separated from the noise for overlapping spectral components where $\Omega r_0/c_0 \gg 1$. In this regime, the signal and noise both satisfy the same plane wave relationships; hence, the difficulty in separating the signal and noise is an obvious result.

The separation of a pulsed sinusoidal signal from bandlimited white Gaussian noise was also addressed for signal-to-noise ratios of the order of unity. Sample functions of the noise were used to demonstrate that the signal could be recovered from the broadband measurements. Although the case of additive extraneous noise was not explicitly considered, it is clear that the effects of such noise become increasingly important as the bandwidth of the measurement system and signal processor increases. In general, the bandwidth of the overall system should of course be matched to that of the signal if the bandwidth is known *a priori*, or, if the signal is known, matched filter methods can then be used.

ACKNOWLEDGMENTS

This work was supported by the Naval Undersea Warfare Center and the Office of Naval Research. The author thanks Dr. Benjamin Cray of NUWC for his helpful discussions and directing the author to the important contribution of the recently deceased A. Parvelescu which formed the foundation for the present paper.

[1] B. Steinberg, *Principles of Aperture and Array System Design*, 1st ed., New York (Wiley, New York, 1976).
[2] W. S. Burdic, *Underwater Acoustic System Analysis*, 2nd ed. (Prentice-Hall, Englewood Cliffs, NJ, 1991).
[3] J. S. Bendat and A. G. Piersol, *Random Data*, 2nd ed. (Wiley, New York, 1986).
[4] T. B. Gabrielson, D. L. Gardner, and S. L. Garrett, ''A simple neutrally buoyand sensor for direct measurement of particle velocity and intensity in water,'' J. Acoust. Soc. Am. **97**, 2227–2237 (1995).
[5] M. J. Berliner and J. F. Lindberg, *Acoustic Particle Velocity Sensors: Design, Performance and Applications* (AIP Press, New York, 1995).
[6] A. Parvelescu, ''Filters for near field noise,'' J. Acoust. Soc. Am. **33**, 1681 (1961).
[7] J. B. Lord, ''Experiments concerning near-field noise filtering,'' J. Acoust. Soc. Am. **34**, 1997 (1962).
[8] A. Papoulis, *Signal Analysis* (McGraw-Hill, New York, 1977).
[9] R. W. Hamming, *Digital Filters* (Prentice-Hall, Englewood Cliffs, NJ, 1977).
[10] A. Oppenheim and R. W. Schafer, *Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1975).
[11] J. Little and L. Schafer, *Matlab/Signal Processing Toolbox* (The Mathworks, Inc., S. Natick, MA, 1988).

Transfer coupler reciprocity: A new low-frequency coupler-reciprocity technique for the absolute calibration of field hydrophones under full environmental conditions

Joseph F. Zalesak
Naval Undersea Warfare Center, Underwater Sound Reference Division, 1176 Howell Street, Newport, Rhode Island 02841-1708

(Received 22 August 1996; revised 16 December 1998; accepted 31 December 1998)

The reciprocity coupler system in the Low-Frequency Facility (LOFAC) at the Underwater Sound Reference Divsion (USRD) of the Naval Undersea Warfare Center is a closed-chamber environmentally controlled system for the absolute calibration of standard hydrophones at low frequencies. There is a need to extend this capability to field hydrophones that are not specifically designed for use in a reciprocity coupler. The transfer reciprocity method described here is an answer to this need. An error analysis is presented indicating the current state of the existing reciprocity coupler including the transfer reciprocity method. The analysis shows that there is a systematic error which can be eliminated by correcting the sensitivities obtained using the transfer reciprocity method. This correction has a frequency independent component of 0.31 dB and a frequency dependent component of 0.1 dB at 700 Hz and 1.0 dB at 2000 Hz. After correction for the systematic error, the uncertainty in sensitivity is ± 0.25 dB for frequencies below 1000 Hz, growing to ± 0.35 dB for frequencies near 2000 Hz. [S0001-4966(99)02304-8]

PACS numbers: 43.58.Vb, 43.30.Xm, 43.30.Yj [SLE]

INTRODUCTION

The reciprocity coupler system in the Low-Frequency Facility (LOFAC) at the Underwater Sound Reference Division (USRD) of the Naval Undersea Warfare Center (NUWC) is a closed-chamber environmentally controlled system for the absolute calibration of standard hydrophones at low frequencies. The coupler is small with an interior volume of only 125 cm^2 that must be precisely known as a function of pressure and temperature. In general, the hydrophone under calibration is part of the boundary of the reciprocity coupler. Therefore the effect of the hydrophone on the mechanical properties of the coupler must be well understood and the free volume of the coupler with the hydrophone in place must be accurately known. There is a need to extend the reciprocity coupler capability to include field hydrophones that are not specifically designed for use in a reciprocity coupler. The problem here is that neither the compliance of the interior of the coupler nor the free volume of the coupler is accurately known when the field hydrophone is in the coupler. The transfer reciprocity method described here is an answer to this need. This method uses two couplers. The first coupler (the reference coupler) uses two well-designed standard reciprocal transducers. Only one of the transducers is required to be reciprocal. The other transducer is required to be reversible and linear. However, in practice, both transducers are of identical construction and hence reciprocal. Our reference coupler has a third port, normally used for calibrating Navy standard hydrophones. The effect of the standard hydrophone on the volume and compliance of the reference chamber can be accurately determined. Other than to close up the third port of the reference coupler, the standard hydrophone plays no role in the transfer reciprocity measurement. Hence, no electrical measurements are required on the standard hydrophone for transfer reciprocity. After electrical measurements are made on the standard reciprocal transducers in the reference coupler, the transducers are transferred to the second coupler (called the transfer coupler) which also contains the hydrophone under test. Electrical measurements are then repeated on the standard reciprocal transducers to determine the effective reciprocity parameter for the transfer coupler. The calibration of the hydrophone under test then proceeds according to the conventional method for coupler reciprocity.

I. THEORY

The reciprocity calibration technique used at USRD is a continuation of the work by Sims and Henriquez.[1] In this technique, one uses the reciprocity parameter J for a closed acoustic chamber filled with fluid which is given by

$$J = \omega C_A = \frac{\omega V}{\rho c^2}, \tag{1}$$

where ω is the angular frequency, C_A is the acoustic compliance, V is the volume of the fluid, ρ is the density of the fluid, and c is the speed of sound in the fluid. In the traditional reciprocity coupler calibration, an acoustic chamber is fitted with two acoustic transducers (one reciprocal transducer referred to as the transducer T, and one sound source referred to as the projector P) as well as the hydrophone H to be calibrated. A typical reciprocity coupler is shown in Fig. 1. Following the notation in Bobber,[2] the free-field voltage sensitivity M_H (which, at low frequencies, is the same as the pressure sensitivity) of the hydrophone is given by

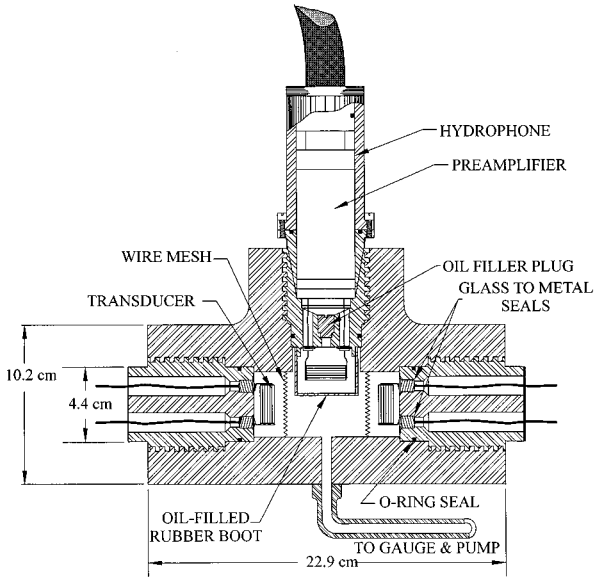


FIG. 1. Typical reciprocity coupler.

$$M_{\mathrm{H}}=\left(J\frac{e_{\mathrm{PH}}}{e_{\mathrm{PT}}}\frac{e_{\mathrm{TH}}}{i_{\mathrm{T}}}\right)^{1/2}, \tag{2}$$

where e_{PH} is the open-circuit output voltage of the hydrophone with the projector operating, e_{TH} is the open-circuit output voltage of the hydrophone with the transducer operating, e_{PT} is the open-circuit output voltage of the transducer with the projector operating, and i_{T} is the electrical current into the transducer. A diagram indicating the electrical inputs and outputs is shown in Fig. 2. A diagram indicating the measurement sequence to perform a conventional reciprocity calibration is shown in Fig. 3.

In practice, the electrical current into the transducer is very small. In order to determine this electrical current, a standard capacitor is placed in series with the transducer and the voltage across the capacitor is measured. If the standard capacitor has a capacitance C, and the voltage across the capacitor is e_{TC}, then

$$i_{\mathrm{T}}=e_{\mathrm{TC}}\omega C. \tag{3}$$

This assumes no leakage across the capacitor and a capacitance independent of frequency. As used here, the standard capacitor behaves as a transducer, converting the electrical

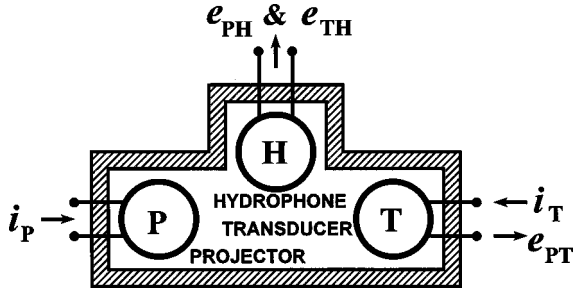


FIG. 2. Diagram of reference reciprocity coupler used for conventional reciprocity. The voltages e_{PH} and e_{PT} are the voltages out of the hydrophone and transducer, respectively, with an electrical current i_{P} into the projector. The voltage e_{TH} is the voltage out of the hydrophone with an electrical current i_{T} into the transducer.

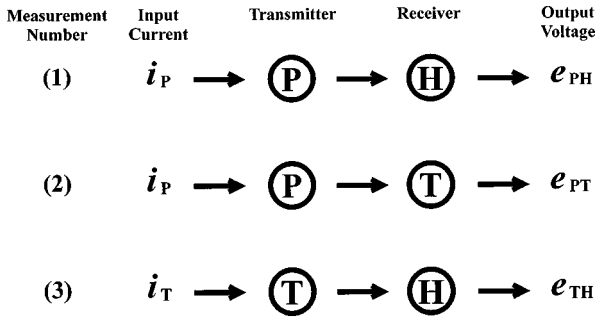


FIG. 3. Diagram of the three measurements for a reciprocity calibration. For the first two measurements the output voltages of the hydrophone e_{PH} and transducer e_{PT} are measured with the same electrical current i_{P} into the projector. For the third measurement, the output voltage e_{TH} of the hydrophone is measured with a measured electrical current i_{T} into the transducer.

current i_{T} into the voltage e_{TC}. By substituting Eqs. (1) and (3) into Eq. (2), one can write the sensitivity of the hydrophone as

$$M_{\mathrm{H}}=\left(\frac{V}{C\rho c^2}\frac{e_{\mathrm{PH}}}{e_{\mathrm{PT}}}\frac{e_{\mathrm{TH}}}{e_{\mathrm{TC}}}\right)^{1/2}. \tag{4}$$

Because of their sizes or their cabling configurations, field hydrophones cannot generally be adapted to fit into the reference reciprocity coupler that we presently use in LOFAC. We have another coupler (called the transfer coupler), originally developed for measurement of the compressibility of elastomeric materials, that is large enough to accommodate field hydrophones and stuffing boxes for the hydrophone cables. However, with the field hydrophone installed, there are uncertainties in the volume and the compliance, and, hence, the reciprocity parameter of the coupler. The presence of a stuffing box in the chamber to bring out the hydrophone cable and the lack of detailed information on the construction and materials of the hydrophone are responsible for the uncertainties. Electrical measurements obtained from both the reference coupler and the transfer coupler are required to do a transfer reciprocity calibration.

The transfer reciprocity technique involves the transfer of reciprocity from a well-characterized or reference coupler such as the one used in LOFAC to a transfer coupler that is less well understood. To accomplish this transfer, electrical measurements are made on the projector and transducer while mounted in the reference coupler. In order to use our reference coupler, the unused port must be closed up. A rigid plug for this port would be ideal. However, since it is a nontrivial task to determine coupler volume with the standard hydrophone and transducers in place, and since the volume measurement has already been made with a standard hydrophone in the third port, we will take advantage of this measurement by using a standard hydrophone in the third port of the reference coupler. No electrical measurements will be made on this standard hydrophone. After the measurements are made on the projector and the transducer in the reference reciprocity coupler, these devices are moved to the transfer coupler for the second part of the transfer reciprocity technique. The transfer coupler must be small enough so that it is operating in its lumped parameter mode, and its projec-

tor must be reversible as well as linear but not necessarily reciprocal. In our system the projector and transducer are both of the same design and, hence, are both linear, reversible, and reciprocal. In this case the roles of projector and transducer can be interchanged, and the additional data obtained can be used to improve the statistics of the measurement or as a check on the data. If we identify subscript 1 with the reference coupler and subscript 2 with the transfer coupler, we can write the two equations

$$e_{\mathrm{TP1}} = M_{\mathrm{P}} p_{\mathrm{TP1}} \tag{5}$$

and

$$e_{\mathrm{TP2}} = M_{\mathrm{P}} p_{\mathrm{TP2}}, \tag{6}$$

where M_{P} is the receiving sensitivity of the projector used as a hydrophone and p_{TP} is the acoustic pressure produced by the transducer at the location of the projector. The assumption made here is that the pressure sensitivity of the projector is equal to its free-field receiving voltage sensitivity. No assumptions are made here about reciprocity of the projector, only reversibility. Since the transducer T is reciprocal, the pressures p at the location of the projector produced in each coupler by the transducer are given by

$$p_{\mathrm{TP1}} = i_{\mathrm{T1}} M_{\mathrm{T}} / J_1 \tag{7}$$

and

$$p_{\mathrm{TP2}} = i_{\mathrm{T2}} M_{\mathrm{T}} / J_2, \tag{8}$$

where i_{T1} and i_{T2} are the electrical currents into the transducer when mounted into the reference coupler and the transfer coupler, respectively, M_{T} is the free-field voltage sensitivity of the transducer, and J_1 and J_2 are the reciprocity parameters for the respective couplers. Substituting Eq. (7) into Eq. (5) and Eq. (8) into Eq. (6) yields

$$e_{\mathrm{TP1}} = M_{\mathrm{P}} i_{\mathrm{T1}} M_{\mathrm{T}} / J_1 \tag{9}$$

and

$$e_{\mathrm{TP2}} = M_{\mathrm{P}} i_{\mathrm{T2}} M_{\mathrm{T}} / J_2. \tag{10}$$

Dividing Eq. (9) by Eq. (10) and simplifying yields

$$\frac{J_1}{J_2} = \frac{i_{\mathrm{T1}} / e_{\mathrm{TP1}}}{i_{\mathrm{T2}} / e_{\mathrm{TP2}}}. \tag{11}$$

Thus we can see that the reciprocity parameter for a coupler is proportional to the transfer admittance of that coupler. Now, substituting the relations

$$i_{\mathrm{T1}} = \omega e_{\mathrm{TC1}} C_1 \tag{12}$$

and

$$i_{\mathrm{T2}} = \omega e_{\mathrm{TC2}} C_2, \tag{13}$$

where C_1 and C_2 are the capacitors used to measure the current when using chambers 1 and 2 (respectively), into Eq. (11) yields

$$\frac{J_1}{J_2} = \frac{e_{\mathrm{TC1}} C_1 / e_{\mathrm{TP1}}}{e_{\mathrm{TC2}} C_2 / e_{\mathrm{TP2}}} \tag{14}$$

or

$$J_2 = J_1 \left(\frac{C_2}{C_1} \frac{e_{\mathrm{TC2}}}{e_{\mathrm{TC1}}} \frac{e_{\mathrm{TP1}}}{e_{\mathrm{TP2}}} \right). \tag{15}$$

If a calibration is made in the transfer coupler we can write the expression for the sensitivity as

$$M_{\mathrm{H}} = \left(\frac{J_2}{\omega C_2} \frac{e_{\mathrm{PH2}}}{e_{\mathrm{PT2}}} \frac{e_{\mathrm{TH2}}}{e_{\mathrm{TC2}}} \right)^{1/2}, \tag{16}$$

by combining Eq. (2) with Eq. (3). If Eq. (15) is substituted into the Eq. (16), and the result simplified using

$$J_1 = \omega V_1 / \rho_1 c_1^2, \tag{17}$$

we obtain

$$M_{\mathrm{H}} = \left[\frac{V_1}{C_1 \rho_1 c_1^2} \frac{e_{\mathrm{TP1}}}{e_{\mathrm{TC1}}} \frac{e_{\mathrm{TH2}}}{e_{\mathrm{TP2}}} \frac{e_{\mathrm{PH2}}}{e_{\mathrm{PT2}}} \right]^{1/2}. \tag{18}$$

None of the fluid parameters of the transfer coupler appear in Eq. (18). Therefore we are free to use different fluids in the reference coupler and the transfer coupler. The acoustic properties of the reference chamber must be accurately known in order to perform an accurate calibration. Other than knowing the acoustic properties of the reference chamber, one must only determine three voltage ratios in order to perform the calibration. One does not have to know the currents into the projector or transducer while performing the measurements in the transfer coupler. Only two voltage ratios are calculated from data obtained from the transfer coupler. The two receiving elements whose voltage outputs are used in forming each of the last two ratios in Eq. (18) must be in identical sound fields when forming the respective ratios. Equation (18) is arranged to show the required ratios to be measured. On examination of the final data, the total random error in the whole expression is approximately 0.01 dB. Thus the random error is not the issue here. The dominating errors are systematic and it is these systematic errors that we must evaluate in order to determine the accuracy of the calculated sensitivity. This systematic error is a function of frequency, i.e., the dominating systematic error will differ in different frequency ranges.

The required steps to perform the transfer reciprocity measurement corresponding to Eq. (18) are shown in Figs. 4, 5, and 6. Figure 4 shows the reference reciprocity coupler set up for transfer reciprocity measurements and Fig. 5 shows the transfer reciprocity coupler set up with the unit under test (UUT) in place. The sequence of six measurements is shown in Fig. 6. Notice that the measurements are grouped in pairs. In each pair of measurements, one transmitter is driven with an electrical current and two voltages are measured to form one of the ratios in Eq. (18). The voltage across a capacitor is used as a measure of the electrical current in measurement (1b) of Fig. 6. In this instance the capacitor itself acts as a receiver in that it converts the desired signal (the electrical current) into a voltage for measurement. The first two measurements of Fig. 6 are made in the reference reciprocity coupler. With an electrical current i_{T1} into the transducer, the voltage e_{TP1} out of the projector and voltage e_{TC1} across the current measurement capacitor are measured. The remainder of the measurements of Fig. 6 are made in the transfer reciprocity coupler. For the third and fourth measurements, the

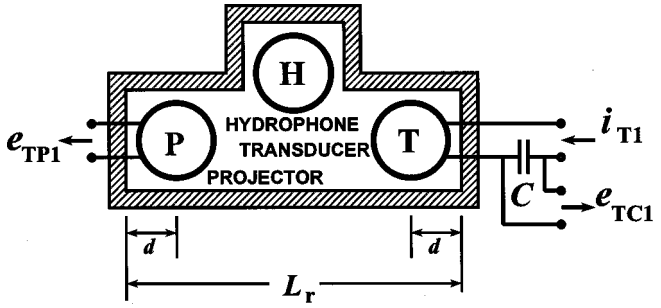


FIG. 4. Diagram of the reference reciprocity coupler used for transfer reciprocity. The acoustic centers of the projector and transducer are located a distance d from the ends of the reference reciprocity coupler. The interior length of the reference reciprocity coupler is L_r. A capacitor C in series with the transducer T is used to measure the electrical current into the transducer. The voltages e_{TP1} and e_{TC1} are the voltage out of the projector and the voltage across the capacitor respectively with an electrical current i_T into the transducer.

voltage e_{PH2} out of the unit under test and the voltage e_{PT2} out of the transducer are measured with an electrical current i_{P2} into the projector. For the fifth and sixth measurements, the voltage e_{TH2} out of the unit under test and the voltage e_{TP2} out of the projector are measured with an electrical current i_{T2} into the transducer.

In our case, the projector is reciprocal; we can interchange the roles of projector and transducer to obtain another measurement of the sensitivity of the field hydrophone. While we would normally take the average of these two measurements, we choose to take the geometric mean because it produces some computational simplicity without sacrificing accuracy (provided that the two values are sufficiently close together). Taking the geometric mean of the two measurements yields

$$M_H = \left[\frac{V_1}{C_1 \rho_1 c_1^2} \left(\frac{e_{TP1}}{e_{TC1}} \frac{e_{PT1}}{e_{PC1}} \right)^{1/2} \frac{e_{TH2}}{e_{TP2}} \frac{e_{PH2}}{e_{PT2}} \right]^{1/2}. \tag{19}$$

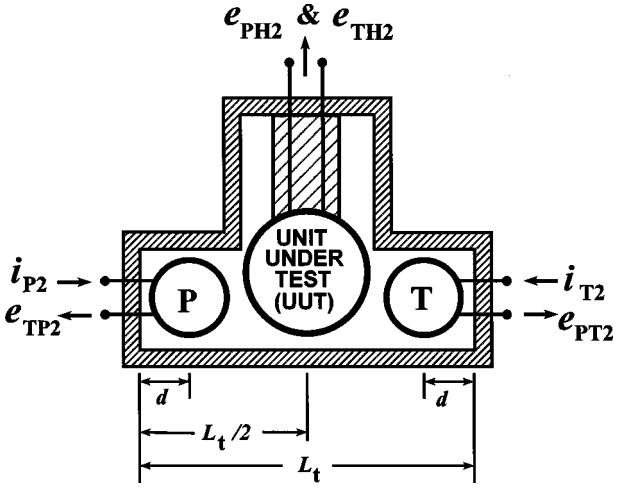


FIG. 5. Diagram of transfer reciprocity coupler with unit under test (UUT) in the hydrophone location. The acoustic centers of the projector and transducer are located a distance d from the ends of the reference reciprocity coupler. The interior length of the transfer reciprocity coupler is L_t. The voltages e_{PT2} and e_{PH2} are the voltages out of the transducer and the unit under test respectively with an electrical current i_{P2} into the projector. The voltages e_{TP2} and e_{TH2} are the voltages out of the projector and the unit under test, respectively, with an electrical current i_{T2} into the transducer.

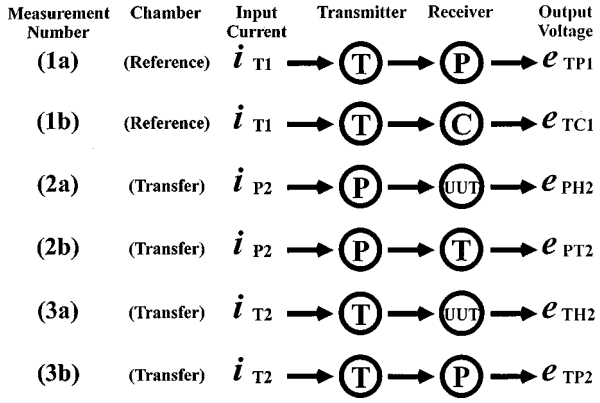


FIG. 6. Diagram of the six measurements for a transfer reciprocity calibration. The measurements are made in pairs, each pair forming one of the ratios in the hydrophone sensitivity formula for transfer reciprocity calibrations. The first two measurements are made in the reference reciprocity coupler. With an electrical current i_{T1} into the transducer, the voltage e_{TP1} out of the projector and voltage e_{TC1} across the current measurement capacitor are measured. The remainder of the measurements are made in the transfer reciprocity coupler. For the third and fourth measurements, the voltage e_{PH2} out of the unit under test and the voltage e_{PT2} out of the transducer are measured with an electrical current i_{P2} into the projector. For the fifth and sixth measurements, the voltage e_{TH2} out of the unit under test and the voltage e_{TP2} out of the projector are measured with an electrical current i_{T2} into the transducer.

II. ERROR ANALYSIS

Since the random errors in the measurement are negligible compared to the systematic errors, we will now proceed to estimate the systematic errors, and the uncertainties in the systematic errors. Any systematic error that is known and that is stable can be estimated and corrected for so that only the uncertainties in the systematic errors remain.

A. Acoustic compliance uncertainty due to uncertainty in fundamental constants, coupler volume measurement, and environmental instability

First we will consider the compliance uncertainty in the reference coupler when filled with castor oil. The quantity $(V_1/\rho_1 c_1^2)$ represents the acoustic compliance of the reference coupler. The volume V_1 of the castor oil in the reference coupler with three transducers in place was previously measured as part of the measurement facility operations. This volume was measured to be $1.247\times10^{-4}\,\mathrm{m}^3$ within an estimated accuracy of $\pm1\%$. Even though only two transducers are required for the transfer reciprocity, the chamber is designed for three transducers and all three must be in place to mechanically complete the chamber for calibration purposes. The calibration was performed at a pressure of 1.793 $\pm$.068 MPa (260$\pm$10 psi) and at a temperature of 4 °C $\pm$1 °C. Under these conditions the nominal density is 971.9 kg/m^3, and the nominal speed of sound is 1563 m/s.[3] The quantity $1/\rho_1 c_1^2$ evaluates to $4.213\times10^{-10}\,\mathrm{m\cdot s^2/kg}\pm0.5\%$ over the pressure and temperature variations indicated above. The dominant contribution to the uncertainty arises from the pressure variation. However, it is the inability to stabilize the temperature of the chamber that causes the pressure variation. An additional uncertainty is due to the inability of the empirically derived formula to accurately predict the speed

of sound in castor oil. The estimated accuracy of the sound speed in castor oil as predicted by this formula is $\pm 1.1\%$.[3] Since the speed of sound appears squared in the expression for the acoustic compliance, the uncertainty in the compliance as a result of the uncertainty in the speed of sound is $\pm 2.2\%$. The uncertainty in the calculated density of castor oil is negligible in relation the other uncertainties and is ignored. These uncertainties add to yield a coupler compliance of $5.254\times 10^{-14}\,\mathrm{m}^3/\mathrm{Pa}$ with an uncertainty of $\pm 3.7\%$. Since the compliance appears under the square root sign in Eq. (18), the resultant uncertainty in hydrophone sensitivity is $\pm 1.9\%$.

B. Corrections and uncertainties as a result of the inability to achieve ideal boundary conditions in reference coupler

1. Reference coupler transducer compliance

It was assumed above that the transducers in the reference coupler are rigid and do not contribute to the compliance. In fact, they are not rigid and add a small contribution that can be calculated using an equation from Roark[4] developed for a hollow sphere subjected to external pressure. The transducers are constructed from Navy Type I lead-zirconate-titanate material. The outside radius is 12.7 mm and the inside radius is 9.53 mm. In this case Young's modulus is $6.45\times 10^{10}\,\mathrm{N/m}^2$ and Poisson's ratio is 0.343 as found in the manufacturer's product data sheet. Using these values, the acoustic compliance is found to be $4.12\times 10^{-16}\,\mathrm{m}^3/\mathrm{Pa}$ which, for the three transducers in the chamber, gives an additional compliance of $1.236\times 10^{-15}\,\mathrm{m}^3/\mathrm{Pa}$. This compliance is a systematic correction which must be added to the castor oil compliance. Application of this correction to the compliance is equivalent to adding 0.10 dB to the sensitivity calculated from Eq. (18) expressed in decibels. Since this is a small correction, the uncertainty in this correction does not add significantly to the overall uncertainty in the measurement.

2. Coupler wall compliance

Another assumption was that the walls of the coupler are rigid, and thus do not contribute to the compliance of the reference coupler. Because of the complicated geometry, it is not easy to accurately calculate this compliance. However, we can obtain an approximation for the compliance as follows. Assume a cylindrical chamber whose thickness is equal to the mean coupler wall thickness. Also assume that the walls at the ends of the cylinder do not bend but do communicate force to the cylindrical portion of the chamber. Using an equation from Roark,[4] one can determine the change in dimensions of such a chamber under uniform internal pressure. In the present case the chamber material is stainless steel type 304. The elastic properties for this material are in standard engineering handbooks.[5] In MKS units, Young's modulus is $1.903\times 10^{11}\,\mathrm{N/m}^2$ and Poisson's ratio is 0.305. For the moment ignore the side arm of the ''tee'' configuration of the reference chamber and only concentrate on the main cylindrical portion of the coupler. The inside radius is 22.23 mm, the mean outside radius is 61.32 mm, and the length is 86.77 mm. The approximate compliance of the chamber walls is $2.252\times 10^{-15}\,\mathrm{m}^3/\mathrm{Pa}$. However, the chamber is not a simple cylinder, but rather has the shape of a ''tee.'' To account for the side arm, one must multiply the above value by the ratio of the total volume to the volume of the main cylinder to get a total contribution of the wall motion to the compliance. The main cylindrical volume is $1.347\times 10^{-4}\,\mathrm{m}^3$ and the total volume is $1.530\times 10^{-4}\,\mathrm{m}^3$. (Note, this volume is not the volume of castor oil in the coupler.) The transducers themselves occupy some of the interior volume of the coupler. The result of applying the volume correction is to yield an approximate wall compliance of $2.559\times 10^{-15}\,\mathrm{m}^3/\mathrm{Pa}$. If we knew this value precisely, we could correct for it in the sensitivity calculation. An independent calculation of the compliance of the reference coupler was performed using finite element analysis.[6] If the results of this analysis are corrected for small differences in geometry, the finite element analysis predicts a chamber compliance of $2.658\times 10^{-15}\,\mathrm{m}^3/\mathrm{Pa}$. Some of the simplifying assumptions used in the finite element analysis cause some uncertainty in the calculated compliance. The error is a result of the inability to model the threaded components of the chamber. Insight into the relative contributions of differing boundary conditions to the total chamber compliance can be obtained by running the finite element code for various boundary conditions and comparing the resultant compliances.[6] In this manner the error is estimated to be $\pm 10\%$ or $\pm 2.66\times 10^{-16}\,\mathrm{m}^3/\mathrm{Pa}$.

3. Total compliance correction

If we add the two compliance corrections above to the compliance of the castor oil, we obtain a total compliance of $5.643\times 10^{-14}\,\mathrm{m}^3/\mathrm{Pa}$ compared with a castor oil compliance of $5.254\times 10^{-14}\,\mathrm{m}^3/\mathrm{Pa}$. Since the compliance $(V_1/\rho_1 c_1^2)$ in Eq. (18) appears under the square root, the correct sensitivity is obtained by multiplying Eq. (18) by the square root of the ratio of the total compliance to the castor oil compliance. The square root of the ratio of the compliances is 1.036. Equivalently, the sensitivities should be increased by 0.31 dB because of this compliance error. The uncertainty in the chamber wall compliance is the dominating uncertainty and is equal to $\pm 2.66\times 10^{-16}\,\mathrm{m}^3/\mathrm{Pa}$. This produces an additional uncertainty of $\pm 0.2\%$ in the calibration predicted by Eq. (18).

C. Electrical current measurement error

Another error to consider is the determination of e_{TC1}. In order to measure the current into the projector, a 1-μF capacitor is placed in series with the projector. The voltage across the 1-μF capacitor is then measured using a differential amplifier. The measured projector capacitance is 4.32 nF. The desired signal is, therefore, 47.3 dB below the common mode signal which must be rejected in order to have a measurement of the voltage across the 1-μF capacitor. The common mode rejection of the differential amplifier is 83 dB as determined from the manufacturer's product literature. Therefore, the theoretical accuracy of the measurement is $\pm 1.6\%$. Since the current in Eq. (18) appears under the

square root sign, the effect on the sensitivity is $\pm 0.8\%$. There is also a small systematic error in the current measurement caused by the conductance of the capacitor and by the frequency dependence of the capacitance of the capacitor. The total impedance of the capacitor should be used in Eq. (18) instead of just the capacitive reactance to determine the current. This error is about 0.012 dB at 50 Hz, and it rapidly decreases to a negligible value at higher frequencies. The correction would decrease the sensitivity from that predicted by Eq. (18). To correct for this, multiply Eq. (18) by the factor K given by

$$K=\left[1+\left(\frac{G_1}{\omega C_1}\right)^2\right]^{-1/4}, \tag{20}$$

where G_1 is the conductance of the capacitor.

D. Acoustic pressure nonuniformity throughout couplers due to finite wavelength effects

Except at extremely low frequencies, there will be a nonnegligible variation in the acoustic pressure throughout the coupler that can affect the calibration. The reference coupler behaves to first order as an acoustic waveguide containing a standing wave driven by a source near one end. Figure 4 shows the geometry. In this approximation we consider the coupler to be a short waveguide whose measured length L_r is given by $L_r=86.77$ mm. The measured location of the hydrophone (in this case T) is a distance $d=14.85$ mm from one end of the coupler. The source P is symmetrically located a distance $d=14.85$ mm from the other end of the coupler. The same P and T are used in both couplers and, hence, the offset distance $d=14.85$ mm is the same in both couplers. The source injects a volume velocity to maintain the field in the chamber. For this approximation the couplers are considered to be waveguides of constant cross section. The projector and transducer are considered to be acoustically rigid. The first order effect of the compliance of these elements was already taken into account in Sec. I above. The term in Eq. (18) being corrected is $(e_{\mathrm{TP1}}/e_{\mathrm{TC1}})$. An acoustic pressure in the reference coupler is responsible for producing the voltage e_{PT1}. The assumed boundary conditions are that the volume velocity is zero at $x=0$ and $x=L_r$, there is a source of volume velocity at $x=(L_r-d)$, the pressure is continuous at the source of volume velocity, and the sum of the volume velocities away from the source is equal to the injected volume velocity. As is shown in the Appendix, the acoustic pressure at the hydrophone location, $x=d$, as a result of wavelength effects, is too high, and as a result, the voltage e_{TP1} is too high by a factor F_1 given by

$$F_1=\cos^2(kd)\left[\frac{kL_r}{\sin(kL_r)}\right]. \tag{21}$$

The transfer coupler is larger than the reference coupler, resulting in a larger wavelength effect. This magnitude of this effect can be derived in a manner similar to that given in the Appendix. The transfer coupler has a measured length $L_t=136.00$ mm with a similar geometry and boundary conditions as the reference coupler with the exception that the unit under test (UUT) is located at the center of the coupler, i.e., at $x=68.00$ mm. The last two terms, $(e_{\mathrm{TH2}}/e_{\mathrm{TP2}})$ and $(e_{\mathrm{PH2}}/e_{\mathrm{PT2}})$, of Eq. (18) are being corrected. In each of these ratios, it was assumed that the pressure is uniform everywhere the chamber. However, the pressure in the coupler is proportional to $\cos(kx)$ where x is the distance from the rigid termination of the coupler. Therefore, the measured ratio is too low by a factor F_2 given by

$$F_2=\frac{\cos(kL_t/2)}{\cos(kd)}. \tag{22}$$

The correction factor F which when applied to Eq. (18) accounts for these wavelength effects is

$$F=\frac{\sqrt{\sin(kL_r)/kL_r}}{\cos(kL_t/2)}. \tag{23}$$

The perhaps not too surprising thing about Eq. (23) is that there is no dependence of the total correction factor on the offset distance d of the acoustic centers of T and P from the rigid ends of the couplers. This offset is fixed by the geometry of P and T causing the error to be identical in the two chambers. The error, thus, cancels. At 2000 Hz, Eq. (23) predicts a correction factor of $+1.0$ dB whereas the droop in the average sensitivity of three field hydrophones at 2000 Hz (a frequency at which the sensitivity is still expected to have the low frequency value) was 0.9 dB. This implies that the overall frequency dependent correction factor has an accuracy of ± 0.1 dB at 2000 Hz. Below 1000 Hz the uncertainty in the correction factor is expected to be negligible.

E. Overall uncertainty and correction factor

A correction of

$$0.31\ \mathrm{dB}+20\log_{10}\left[\frac{\sqrt{(\sin(kL_r)/kL_r)}}{\cos(kL_t/2)}\right] \tag{24}$$

should be added to all sensitivities obtained from using Eq. (17). The frequency dependent portion of the correction is 0.1 dB at 700 Hz and 1.0 dB at 2000 Hz. After application of the above correction, the uncertainty in the measurement including the effects of compliance uncertainty and current uncertainty is $\pm 2.9\%$ or ± 0.25 dB for frequencies up to 1000 Hz. For frequencies near 2000 Hz the uncertainty increases to about ± 0.35 dB because of the additional uncertainty in the frequency dependent correction factor. Three similarly constructed hydrophones were calibrated using the transfer reciprocity method. These hydrophones are expected to have a constant sensitivity in the frequency range of 100 Hz–2000 Hz. Since the absolute sensitivities of these hydrophones are not expected to be identical, each of their sensitivities was normalized to 0.0 dB over the frequency range of 100 Hz–1000 Hz for ease in comparison. The normalized sensitivities are shown in Fig. 7. Note how the sensitivities roll off at the higher frequencies. Also note the consistency in the shape of the sensitivity curves. These sensitivities were then corrected using Eq. (24). Each individual curve was again normalized to have an average sensitivity of 0.0 dB over the frequency range of 100 Hz–1000 Hz. The resulting curves are shown in Fig. 8. At the higher frequencies the corrected sensitivities are increasing with frequency indicating that the curves have

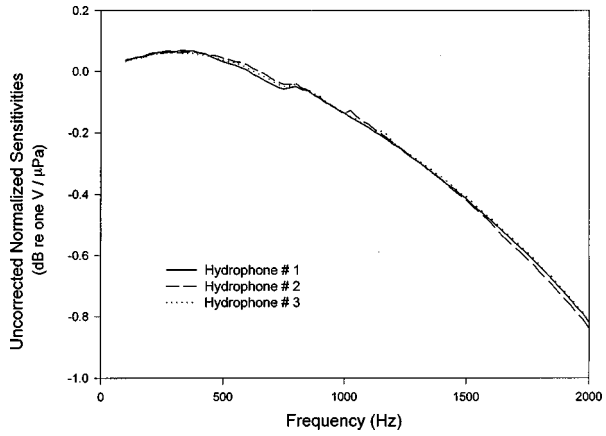


FIG. 7. Sensitivities of three similarly constructed hydrophones measured using the transfer reciprocity method and normalized to have an average sensitivity of 0.0 dB over the frequency range 100 Hz–1000 Hz.

been slightly over corrected. Notice, again, the consistency of the shape of the sensitivity curves. On this scale, the random error is just becoming evident. The fine structure near 700 Hz and the slight roll off for frequencies below 500 Hz are unexplained. Correcting the sensitivity for the conductance in the current measurement capacitor will slightly increase the slope of the curve for frequencies below 100 Hz.

III. CONCLUSIONS

Our present transfer reciprocity system has a systematic error of

$$0.31\ \mathrm{dB} + 20\log_{10}\left[\frac{\sqrt{(\sin(kL_r)/kL_r)}}{\cos(kL_t/2)}\right] \tag{25}$$

which must be added to the calculated sensitivity. The uncertainty after correcting for the systematic error is ± 0.25 dB for frequencies up to 1000 Hz, increasing to about ± 0.35 dB for frequencies near 2000 Hz. There is a slight unexplained roll off in sensitivity for frequencies below 500 Hz. Correcting for the conductance in the current measurement capacitor increases the roll off slightly for frequencies below 100 Hz.

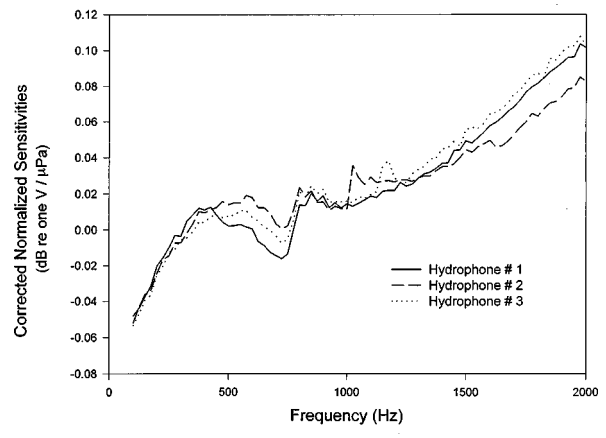


FIG. 8. Corrected sensitivities of three similarly constructed hydrophones measured using the transfer reciprocity method and normalized to have an average sensitivity of 0.0 dB over the frequency range 100 Hz–1000 Hz.

IV. RECOMMENDATIONS FOR IMPROVING ACCURACY

Improve the measurement of the electrical current so that the uncertainty in the electrical current is negligible. Modify the reference coupler geometry in order to minimize the wall compliance and to be better able to calculate the wall compliance. Replace the projector and transducer with elements whose compliance is much smaller so that their contributions to the overall compliance is reduced. Use deionized water in the reference coupler to reduce the uncertainty in the compliance of the fluid so that its uncertainty does not dominate the error budget. Improve temperature stability of the coupler to reduce uncertainty because of temperature drift. Reduce overall dimensions of both couplers to minimize pressure nonuniformity effects as a function of frequency. Design the couplers so that elements used as hydrophones are symmetrically located with respect to the sound source and, hence, experience an equivalent acoustic field.

Stability of P and T as hydrophones is required for the transfer reciprocity technique. They are, in fact, being used as reference hydrophones in the transfer coupler. As a general rule, one should not use reference hydrophones as sound sources because of the danger of the high applied electric fields causing sensitivity changes primarily as a result of accelerated aging and possibly depolarization. To see that it is not necessary to use T as a sound source in the transfer coupler note the term (e_{TH2}/e_{TP2}) in either Eq. (18) or Eq. (19). This term represents the outputs of H and P in the transfer coupler using T for the sound source. Note that neither Eq. (18) nor Eq. (19) require any data from T itself. The ratio (e_{TH2}/e_{TP2}) is simply the ratio of two voltages obtained from the two hydrophones when they are ensonified equally. The same ratio would be obtained using any source that ensonifies H and P equally. Similarly, the measurement of the ratio (e_{PH2}/e_{PT2}) does not require P to be the sound source. The only requirement is that H and T be ensonified equally. Since neither T nor P are required to be sound sources in the transfer chamber, it is recommended that separate sound source be used in the transfer coupler. The transfer coupler would contain four acoustic devices; H, T, P, and the newly identified sound source. In this manner, P and T would only be used as hydrophones in the transfer coupler thus minimizing the possibility that the sensitivities of P and T will change over time.

ACKNOWLEDGMENT

This work was supported by the Office of Naval Research via the Naval Surface Warfare Center, Carderock Division.

APPENDIX: DERIVATION OF WAVELENGTH DEPENDENT CORRECTION FACTOR EQ. 21

We show derivation of Eq. (21) using the velocity potential for plane waves.[7] The relative positions of the transducer and projector with respect to the ends of the coupler are shown in Fig 4. Assume that the origin of the coordinate system is at the left end of the interior volume of the coupler. Also assume a waveguide of constant cross section. Assume

that the transducer injects a volume velocity U into the waveguide. The velocity potential ϕ for plane waves traveling along the x-direction is

$$\phi=[A\cos(kx)+B\sin(kx)]e^{j\omega t}, \tag{A1}$$

where A and B are complex constants to be determined from the boundary conditions, $\omega=2\pi f$, f is the frequency of operation, $k=\omega/c$, c is the velocity of sound in the fluid, and t is time. The acoustic pressure p and the particle velocity u are determined from the velocity potential by

$$p=\rho\frac{\partial\phi}{\partial t}=j\omega\rho[A\cos(kx)+B\sin(kx)]e^{j\omega t}, \tag{A2}$$

and

$$u=-\frac{\partial\phi}{\partial x}=k[A\sin(kx)-B\cos(kx)]e^{j\omega t}. \tag{A3}$$

The explicit time dependence will not be shown in the equations that follow. The volume velocity U is equal to the product of the particle velocity u and the cross section S of the coupler, i.e.,

$$U=Su. \tag{A4}$$

The coupler has a volume velocity source located at $x=L_r-d$ as shown in Fig. 4. This source will be approximated by a uniform, planar, infinitesimally thin velocity source spanning the cross section at $x=L_r-d$. The velocity potential is only valid in the region where there are no sources. In this case there is a velocity potential ϕ_L valid to the left of the source and another velocity potential ϕ_R valid to the right of the source, i.e.,

$$\phi_L=A_L\cos(kx)+B_L\sin(kx), \tag{A5}$$

and

$$\phi_R=A_R\cos(kx)+B_R\sin(kx). \tag{A6}$$

The corresponding pressures and volume velocities are given by

$$P_L=j\omega\rho[A_L\cos(kx)+B_L\sin(kx)], \tag{A7}$$

$$P_R=j\omega\rho[A_R\cos(kx)+B_R\sin(kx)], \tag{A8}$$

$$U_L=kS[A_L\sin(kx)-B_L\cos(kx)], \tag{A9}$$

and

$$U_R=kS[A_R\sin(kx)-B_R\cos(kx)]. \tag{A10}$$

Applying the boundary condition $U_L=0$ at $x=0$ to Eq. (A9) yields

$$B_L=0. \tag{A11}$$

Applying the boundary condition $U_R=0$ at $x=L_r$ to Eq. (A10) yields

$$B_R=A_R\tan(kL_r). \tag{A12}$$

Applying continuity of pressure at the location of the source, i.e., $p_L=p_R$ at $x=L_r-d$, to Eqs. (A7) and (A8) and simplifying using Eqs. (A11) and (A12) yields

$$A_R=A_L\frac{\cos(kL_r)\cos(kL_r-kd)}{\cos(kd)}. \tag{A13}$$

Applying the condition that at the location of the source the injected volume velocity is equal to the sum of the volume velocities directed away from the source, i.e., $U=U_R-U_L$ at $x=L_r-d$ to Eqs. (A7) and (A8) and simplifying using Eqs. (A11)–(A13) yields

$$A_L=-\frac{U}{kS}\frac{\cos(kd)}{\cos(kL_r)}. \tag{A14}$$

Substituting Eq. (A14) into Eq. (A13) yields

$$A_R=-\frac{U}{kS}\frac{\cos(kL_r)\cos(kL_r-kd)}{\sin(kL_r)}. \tag{A15}$$

Substituting Eq. (A15) into Eq. (A12) yields

$$B_R=-\frac{U}{kS}\cos(kL_r-kd). \tag{A16}$$

Thus all the constants in the velocity potentials have been solved for. At low frequencies, the pressure p_{lf} can be determined from Eq. (A7) by limiting ω to small values. The result is

$$p_{lf}=\frac{U}{j\omega(V/\rho c^2)}, \tag{A17}$$

where the volume of the coupler $V=L_rS$ and the quantity $(V/\rho c^2)$ is the acoustic compliance of the coupler. The ratio of the pressure at $x=d$ to the low frequency pressure is the factor F_1 of Eq. (21). An expression for the factor F_1 is obtained by taking the ratio of Eq. (A7) evaluated at $x=d$ and Eq. (A17). The result is

$$F_1=\frac{p_L(x=d)}{p_{lf}}=\frac{kL_r\cos^2(kd)}{\sin(kL_r)}. \tag{A18}$$

This wavelength correction is a small correction so it is not necessary to know it precisely. Attempts to further improve this calculation are not worth pursuing because of the complexity of the calculations and the uncertainties in the true boundary conditions imposed by all the items within the couplers and the configurations of the couplers themselves.

[1] C. C. Sims and T. A. Henriquez, ''Reciprocity calibration of a standard hydrophone at 1600 psi,'' J. Acoust. Soc. Am. **36**, 1704–1707 (1964).

[2] R. J. Bobber, *Underwater Electroacoustic Measurements* (Peninsula, Los Altos, CA, 1988).

[3] R. W. Timme, ''Speed of sound in castor oil,'' J. Acoust. Soc. Am. **52**, 989–992 (1972).

[4] R. J. Roark and W. C. Young, *Formulas for Stress and Strain*, 5th ed. (McGraw-Hill, New York, 1982).

[5] T. Baumeister, E. Aballone, and T. Baumiester III, *Marks' Standard Handbook for Mechanical Engineers*, 8th ed. (McGraw-Hill, New York, 1978).

[6] S. C. Black and C. M. Siders, ''Finite element modeling for the low frequency facility's reciprocity coupler,'' Naval Research Laboratory, Underwater Sound Reference Detachment, internal project report (unpublished), January 12, 1995.

[7] R. W. B. Stephens and A. E. Bate, *Acoustics and Vibrational Physics*, 2nd ed. (Edward Arnold, London, 1966).

Power estimation of sound sources on low-speed electric trains using a deconvolution approach

W. M. To and S. M. Yung
Mechanical Engineering Department, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong

(Received 6 December 1997; revised 28 October 1998; accepted 7 January 1999)

The traditional delay and sum beamforming method has been widely used over the past two decades to locate and characterize sound sources on moving trains. However, when the total number of major and minor sources is uncertain, this method tends to overestimate the sound powers for multiple sources. The overestimation of sound powers becomes significant when the separation of the sources is very ''small'' for a microphone array with fixed configuration for avoiding spatial aliasing at high frequencies for broadband noise. The separation of the sources is ''small'' because the resulting resolution of the delay and sum beamforming method is ''low'' at low frequencies. For train noise characterization, the contribution from some other parts of the train would be masked due to this low resolution power at low frequencies. This paper describes a deconvolution approach which is developed to give estimates of the sound powers on a moving line source. Simulation results illustrate that sound powers can be accurately estimated by using the proposed deconvolution method at both low- and high-frequency regions. Both the delay and sum beamforming method and the deconvolution method were applied to the measured data using an array of 15 microphones. © *1999 Acoustical Society of America.* [S0001-4966(99)02004-4]

PACS numbers: 43.60.Gk, 43.20.Ye, 43.58.Fm [JCB]

INTRODUCTION

It is well known that, with appropriate array signal processing techniques, an array of microphones can provide high resolution on directionality and the ability to identify the location of major sound sources. The delay and sum beamforming method (DAS), which is one of the traditional array signal processing methods mentioned in Johnson and Dudgeon,[1] has been widely used to determine major source locations. Billingsley and Kinns[2] used the DAS method to locate the major sound sources on a jet engine. King III,[3] King III and Bechert,[4] Barsikow *et al.*,[5] and Barsikow[6] modified the DAS method by incorporating a sweeping focus to locate the major sound sources on moving trains. Barsikow *et al.*[5] and Barsikow[6] illustrated that the sweeping focus DAS method can be employed for octave band analysis even though this method is frequency dependent. It is because train noise has strong characteristic peaks at certain frequencies. Other high resolution array signal processing methods, such as MUSIC[7] and ESPRIT,[8] were developed in radar and sonar applications. These methods are all subject to a constraint that the number of sensors must be larger than the number of sources (degrees of freedom). However, the sweeping focus DAS method can be used to localize major sources when the number of sources is larger than the number of sensors.[3–6]

Yung and To[9] constructed an array of microphones to measure noise from moving trains. They employed the sweeping focus DAS method to analyze sound emitted from various parts of the moving trains. They found that the relative contribution of sound energy from various parts of the train could not be accurately estimated because of the masking effect due to the small separation of sound sources. However, they summarized from the measurement results and the experience gained by Barsikow[5] that the major sound sources are located around train wheels for electric trains with speeds lower than 250 km/h. Other sources including traction motors, the rails, and air conditioners could be considered as minor sound sources. For stationary sources at known directions, Lardies[10] proposed a number of unbiased power estimators for multiple sources by using a linear acoustical array. Fuchs *et al.*[11,12] applied a deconvolution approach to the output of the delay and sum beamformer for power estimation. They showed that the deconvolution approach can indeed be seen as a means to obtain maximum likelihood estimates when the beamformer output is taken as the observed data.

Because knowing the location and the relative contribution of the sound sources are important for noise control purposes and for developing new models of quiet trains, an iterative deconvolution method is proposed to give better estimates of the sound powers on a moving line source. Numerical simulations were performed to evaluate the applicability of the proposed method on moving line sources. Both the sweeping focus DAS and the deconvolution methods were then utilized to determine the location and estimate the power of sound sources on moving trains.

I. THEORY

In what follows, scalar values are represented by small/capital letters, vectors are represented by bold small letters, matrices are represented by bold capital letters, superscript ^{T} means to transpose, and $E[\ ]$ means an expected value.

A. Delay and sum beamforming (DAS) method

The basic principle of the DAS method is that the propagation times of sound from sources from various directions

 0001-4966/99/105(4)/2350/9/$15.00 © 1999 Acoustical Society of America

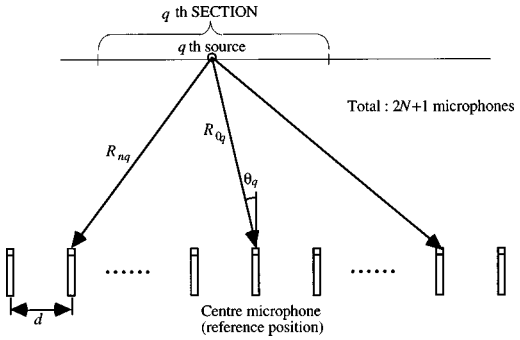


FIG. 1. Array configuration.

to each microphone in the array are different. An array consisting of $2N+1$ microphones with a separation of d meters is shown in Fig. 1. A continuous line source is equally divided into Q sections and each section is represented by a discrete source with a power equal to the total power of that section. The qth discrete source is assumed to be at an angle θ_q, with the array axis and power μ_q.

The array output of the DAS method when steered (focus) to the qth source is

$$z_q(t)=\frac{1}{(2N+1)}\sum_{n=-N}^{N} s_n\left(t-\frac{R_{nq}-R_{0q}}{c}\right), \tag{1}$$

where $s_n(t)$ is the sound pressure received by the nth microphone at time t, R_{nq} is the distance between the qth source and the nth microphone, R_{0q} is the distance between the qth source and the center microphone, and c is the sound speed.

The array output vector is defined as

$$\mathbf{z}_q=[z_q(-T)\, z_q(-T+1)\, \cdots\, z_q(0)\, \cdots\, z_q(t)\, \cdots\, z_q(T)], \tag{2}$$

where $2T+1$ is the number of snapshots.

When no prior information about the source locations is available, the array is steered to a set of sampled directions to cover -90 to 90 degrees, say from $\theta_{-M}, \theta_{-(M-1)}, \ldots, \theta_0, \ldots, \theta_M$, where θ_m is the angle with the array axis. The array output $\mathbf{z}_m$ for each direction can then be computed. The steering response $\mathbf{p}$ is the powers estimated for different directions,

$$\mathbf{p}=\begin{bmatrix} p_{-M} \\ p_{-(M-1)} \\ \cdots \\ p_0 \\ \cdots \\ p_M \end{bmatrix} = \frac{\rho c}{2\pi}\begin{bmatrix} E[z^2_{-M}(t)] \\ E[z^2_{-(M-1)}(t)] \\ \cdots \\ E[z^2_0(t)] \\ \cdots \\ E[z^2_M(t)] \end{bmatrix} = g\begin{bmatrix} E[\mathbf{z}_{-M}\mathbf{z}^T_{-M}] \\ E[\mathbf{z}_{-(M-1)}\mathbf{z}^T_{-(M-1)}] \\ \cdots \\ E[\mathbf{z}_0\mathbf{z}^T_0] \\ \cdots \\ E[\mathbf{z}_M\mathbf{z}^T_M] \end{bmatrix}, \tag{3}$$

where ρ is the density of air and g is a constant converting the estimated powers from the sounds received by the microphones.

For a monochromatic sound source located at an angle θ_q with the array axis, the sound received by the center microphone array is

$$s_0(t)=a_q \exp(j2\pi f_0 t), \tag{4}$$

where a_q is the amplitude of sound pressure and f_0 is the source frequency.

The sound received by the nth microphone is

$$s_n(t)=a_q\frac{R_{0q}}{R_{nq}}\exp\left(j2\pi f_0\left(t+\frac{nd\sin\theta_q}{c}\right)\right). \tag{5}$$

Substituting Eq. (5) into Eq. (1), the array output steered to θ_m becomes

$$z_m(t)=\frac{1}{2N+1}\sum_{n=-N}^{N} a_q\frac{R_{0q}}{R_{Nq}} \times\exp\left(j2\pi f_0\left(t+\frac{nd\sin\theta_q}{c}-\frac{nd\sin\theta_m}{c}\right)\right). \tag{6}$$

When the distance between the sources and the array is much greater than the separation of microphones, Eq. (6) becomes

$$\begin{aligned} z_m(t)&=\frac{1}{2N+1}a_q\exp(j2\pi f_0 t)\sum_{n=-N}^{N}\exp(-jn\beta) \\ &=\frac{1}{2N+1}a_q\exp(j2\pi f_0 t)\exp(jN\beta) \\ &\quad\times\sum_{n=0}^{2N}\exp(-jn\beta), \end{aligned} \tag{7}$$

where $\beta=k_0 d(\sin\theta_m-\sin\theta_q)$ and $k_0=2\pi f_0/c$.

Using the formula for the sum of a geometric series,

$$\sum_{n=0}^{N-1} b_n=\frac{1-b^N}{1-b}, \tag{8}$$

$z_m(t)$ can be written as

$$z_m(t)=\frac{1}{2N+1}a_q\exp(j2\pi f_0 t)\frac{\sin[(2N+1)/2]\beta}{\sin\frac{1}{2}\beta}. \tag{9}$$

Substituting Eq. (9) into Eq. (3), the mth element of the steering response is expressed as

$$p_m=\frac{1}{(2N+1)^2}\mu_q\left[\frac{\sin^2[(2N+1)/2]k_0 d(\sin\theta_m-\sin\theta_q)}{\sin^2\frac{1}{2}k_0 d(\sin\theta_m-\sin\theta_q)}\right], \tag{10}$$

where $\mu_q=gE[a_q^2(t)]$ is the power of the qth source. It should be noted that the steering response is at its maximum when $\theta_m=\theta_q$, and it is possible to locate the source direction from the steering response in this simple case.

The array pattern of steering to direction θ_m when there is a source at θ_q is

$$w_m(\theta_q)=\frac{1}{(2N+1)^2}\left[\frac{\sin^2[(2N+1)/2]k_0 d(\sin\theta_m-\sin\theta_q)}{\sin^2\frac{1}{2}k_0 d(\sin\theta_m-\sin\theta_q)}\right]. \tag{11}$$

The pattern consists of a mainlobe in the focus direction and has several sidelobes. The width of the mainlobe governs the

resolving power of the array. The narrower the mainlobe is, the higher the resolving power is. By increasing the number of microphones, the separation between microphones, and the frequency of interest, the mainlobe width can be reduced. Moreover, the mainlobe becomes narrower when the focus direction θ_m is close to the dead ahead point. Since the array pattern does not have an impulse shape, signals from other directions can pass through the spatial filter and leak energy to the focus direction, resulting in the overestimate of source power.

For a fixed number of microphones and a specific frequency, increasing the separation between microphones may cause spatial aliasing. Spatial aliasing is the phenomenon when another mainlobe appears at the direction other than the focus direction.

B. Sweeping focus delay and sum method

For moving sources, provided that their locations are known at each time instant, the focus can be swept with the same speed as the sources (Bariskow *et al.*[5]) to track the target. In doing so, the Doppler frequency shift effect is eliminated because the observer is moving with the same speed and direction as the sources. When one of the sources passes through the array axis at $t=0$, the angle of the focus with the array axis on that source at any time is

$$\theta_q(t)=\tan^{-1}\frac{Vt}{D}, \tag{12}$$

where D is the shortest distance between that source and the center microphone. As the focus is a time variant, the average array pattern for sweeping focus DAS method is

$$w_m(\theta_q)=\frac{1}{(2N+1)^2}\frac{1}{2T+1}\times\sum_{t=-T}^{T}\left[\frac{\sin^2[(2N+1)/2]k_0d(\sin\theta_m-\sin\theta_q(t))}{\sin^2\frac{1}{2}k_0d(\sin\theta_m-\sin\theta_q(t))}\right], \tag{13}$$

where $2T+1$ is the number of snap shots in tracking.

A row vector of the array pattern $\mathbf{w}_m$ can be constructed using Eq. (13). The array output and the steering response are defined similarly as the case for stationary sources and are denoted as $\mathbf{z}'_q$ and $\mathbf{p}'$.

C. Deconvolution method

Deconvolution is a process of constructing the input from the output and the system response. Lardies[10] and Fuchs *et al.*[11,12] used this approach to determine the locations of multiple stationary sources.

For noncoherent sources, the effect of sources on the steering response is additive. The relation between the source input, the steering response, and the array pattern can be expressed as

$$\begin{bmatrix} p_{-M} \\ p_{-(M-1)} \\ \cdots \\ p_0 \\ \cdots \\ p_M \end{bmatrix}_{(2M+1)\times 1} = \begin{bmatrix} \mathbf{W}'_{-M} \\ \mathbf{W}'_{-M+1} \\ \cdots \\ \mathbf{W}'_0 \\ \cdots \\ \mathbf{W}'_M \end{bmatrix}_{(2M+1)\times(2Q+1)} \times \begin{bmatrix} \mu_{-Q} \\ \mu_{-Q+1} \\ \cdots \\ \mu_0 \\ \cdots \\ \mu_Q \end{bmatrix}_{(2Q+1)\times 1} \tag{14}$$

or

$$\mathbf{p}'=\mathbf{W}'\mathbf{u}'. \tag{15}$$

From noise measurements, one can obtain the vector $\mathbf{p}'$ accurately. The array pattern matrix $\mathbf{W}'$ can be determined from the geometry of the measurement setup. The vector of source powers $\mathbf{u}'$ cannot be measured directly, but its estimate $\hat{\mathbf{u}}$ can be obtained by using the following minimization:

$$\hat{\mathbf{u}}=\min_{\hat{\mathbf{u}}}(\text{norm}\,(\mathbf{p}'-\mathbf{W}'\hat{\mathbf{u}})). \tag{16}$$

D. A new iterative deconvolution scheme

Morse and Ingard[13] studied the sound pressure generated by a moving sound source. If the source strength is

$$q(t)=q_0\sin(\omega_0 t) \tag{17}$$

and the source moving along the x axis passes the origin at $t=0$, the distance between the source and the observer [at (x_0,y_0,z_0)] can be expressed as

$$R=\frac{M_a(x-Vt)+\sqrt{(x-Vt)^2+(1-M_a^2)(y_0^2+z_0^2)}}{1-M_a^2}, \tag{18}$$

where V is the subsonic source speed and M_a is the Mach number ($M_a=V/c$).

The sound pressure received at the observer is

$$p(t)=\frac{q_0\omega_0\cos(\omega_0(t-R/c))}{4\pi R(1-M_a\cos\theta)^2}+\frac{q_0\sin(\omega_0(t-R/c))}{4\pi}\frac{(\cos\theta-M_a)V}{R^2(1-M_a\cos\theta)^3}, \tag{19}$$

where $R_1=\sqrt{(x-Vt)^2+(1-M_a^2)(y_0^2+z_0^2)}$ and $\cos\theta=(R-R_1)/M_aR$.

Now one can define the appearing amplitude of the sound pressure at the observer as

$$a=\frac{q_0}{4\pi R(1-M_a\cos\theta)^2}\sqrt{\omega_0^2+\frac{(\cos\theta-M_a)^2V^2}{R^2(1-M_a\cos\theta)^2}}. \tag{20}$$

It is well known[5] that the major sound sources on a low speed electric train are at the wheel locations. As a result, it is reasonable to assume that there are several discrete major sound sources (at the wheel locations) on the train and that

other parts of the train create by discrete minor sound sources. The velocity of the train can be determined by using a pair of photoelectric sensors.

Assume that a moving line source is equally divided to $2Q+1$ sections and has only two uncorrelated major sources located at θ_{m_1} and θ_{m_2} and with powers μ_{m_1} and μ_{m_2}, respectively. The remaining sources are minor sources and their powers are assumed to be μ_0. Due to the moving effect and the distance attenuation, the amplitude of sound pressure at various receiving positions is different and can be obtained using Eq. (20). The intensity/power factors of the minor and major sources, which are proportional to the square of the receiving amplitudes, are

$$\mathbf{A}_{\text{minor}} = \text{diag}(\overline{a^2}_{-M}\ \overline{a^2}_{-M+1} \cdots \overline{a^2}_{m_1-1}\ 0\ \overline{a^2}_{m_1+1} \cdots \overline{a^2}_{m_2-1}\ 0\ \overline{a^2}_{m_2+1} \cdots \overline{a^2}_{M}) \tag{21}$$

and

$$\mathbf{A}_{\text{major}} = \text{diag}\,(\overline{a^2}_{m_1}\quad \overline{a^2}_{m_2}), \tag{22}$$

where $\overline{a}_m$ is the mean value of the amplitude computed by Eq. (20) during the integration period for the mth source. The vector representing the energy received by microphones is

$$\mathbf{u} = [\overline{a^2}_{-M}\mu_0 \cdots \overline{a^2}_{m_1-1}\mu_0\ \underbrace{\overline{a^2}_{m_1}\mu_{m_1}}_{\text{the } m_1 \text{ element}}\ \overline{a^2}_{m_1+1}\mu_0 \cdots \underbrace{\overline{a^2}_{m_2}\mu_{m_2}}_{\text{the } m_2 \text{ element}} \cdots \overline{a^2}_{M}\mu_0]^T_{(2M+1)\times 1}. \tag{23}$$

A new iterative scheme based on the steering response of the DAS and deconvolution methods is then developed.

Step 1: Generate the steering response of the sweeping focus DAS method, $\mathbf{z}'_m$, for $m=-M$ to M with the focus tracking one of the sources for $t=-T$ to T. Then convolute the steering response with an octave band filter and obtain the filtered steering response to be $\mathbf{p}'_f$.

Step 2: Generate the moving array pattern of the array using Eq. (12) for $t=-T$ to T and the frequency is the center frequency of the filter used above.

Step 3: Assume that the mean power level of minor sources is $\hat{\mu}_b$. Define the vector of minor source power estimates as

$$\hat{\mathbf{u}}_{\text{minor}} = [\hat{\mu}_b \quad \hat{\mu}_b \quad \cdots \quad \hat{\mu}_b]^T_{(2M+1)\times 1}. \tag{24}$$

The steering response due to the minor background sources only can be written as

$$\mathbf{p}'_{\text{minor}} = \mathbf{W}'\,\mathbf{A}_{\text{minor}}\hat{\mathbf{u}}_{\text{minor}}. \tag{25}$$

Step 4: Substract the actual filtered steering response $\mathbf{p}'_f$ by $\mathbf{p}'_{\text{minor}}$ to generate the so-called steering response due to major sources $\mathbf{p}'_{\text{major}}$:

$$\mathbf{p}'_{\text{major}} = \mathbf{p}'_f - \mathbf{p}'_{\text{minor}}. \tag{26}$$

Some elements of $\mathbf{p}'_{\text{major}}$ may be negative, and they are designated as zeros.

Step 5: Define the response of the major sources as

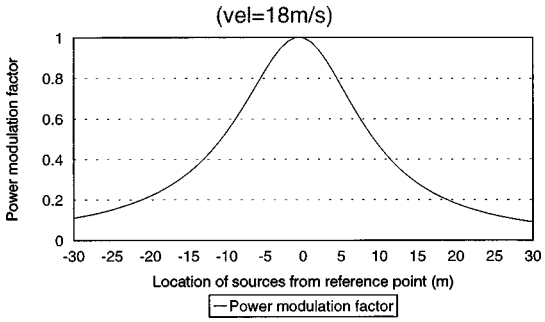


FIG. 2. Power modulation factor of sources at different location.

$$\mathbf{W}'_{\text{major}} = [\mathbf{w}'_{m1} \quad \mathbf{w}'_{m2}] \tag{27}$$

and define the estimate of major sources power vector as

$$\hat{\mathbf{u}}_{\text{major}} = [\mu_{m_1} \quad \mu_{m_2}]^T, \tag{28}$$

where $\hat{\mathbf{u}}_{\text{major}}$ is obtained from the following minimization:

$$\hat{\mathbf{u}}_{\text{major}} = \min_{\hat{\mathbf{u}}_{\text{major}}} (\text{norm}(\mathbf{p}'_{\text{major}} - \mathbf{W}'_{\text{major}}\mathbf{A}_{\text{major}}\hat{\mathbf{u}}_{\text{major}})). \tag{29}$$

Step 6: Define the sum of squared errors between the actual steering response and the steering response from the estimates of the qth iteration as

$$\varepsilon_q = \sum_{m=-M}^{M} (p'_m - \mathbf{w}_m^T\mathbf{A}\hat{\mathbf{u}})^2, \tag{30}$$

where p'_m is the mth element of the actual filtered steering response $\mathbf{p}'_f$.

Step 7: Repeat steps 3–6 for increasing values of $\hat{\mu}_b$. The value of ε_q will vary with $\hat{\mu}_b$ and exhibit a unique minimum. The best estimate of $\hat{\mathbf{u}}$ is the one with the smallest ε_q.

When the length of the line source is extremely long, Q is large and thus the computation time can be prohibitively long. As the number of equations and variables will affect the performance of the power estimation and the computation time, several points have to be considered for determining the optimal value for Q and M. The amplitude modulation due to the distance attenuation and the movement of sound sources will reduce the sound energy received by each microphone when the source is located further away from the dead ahead point (see Fig. 2). Moreover, the average transformed array pattern $\mathbf{w}'_m$ has a very sharp peak when its focus is steered close to the array axis as shown in Fig. 3. Taking these two factors into account, one can conclude that when the focus is steered close to the array axis, the sources located far away from the dead ahead point will have little contribution to the steering response of sweeping focus DAS method. When the array is steered to a wide angle with respect to the dead ahead point, the average transformed array pattern $\mathbf{w}'_m$ has a wide mainlobe as shown in Fig. 3. In considering the shape of the mainlobe and the distance attenuation on source strength, the sources close to the focus may have more or less the same energy contribution as those from the sources close to the dead ahead point. In this situation approximating the output by a limited number of sources

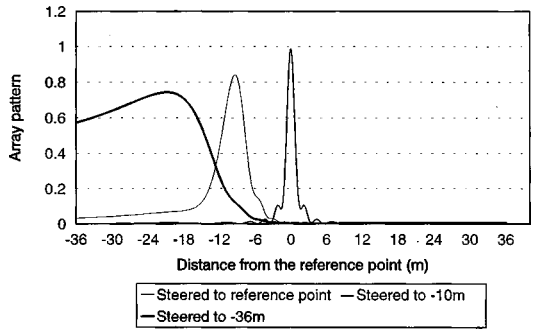


FIG. 3. Array patterns steered to different direction.

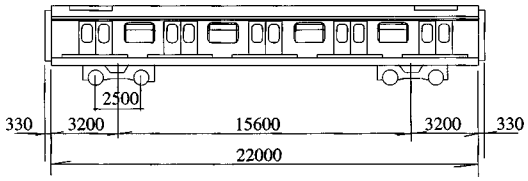


FIG. 4. Detailed drawing of an electric multiple unit (all dimensions in mm).

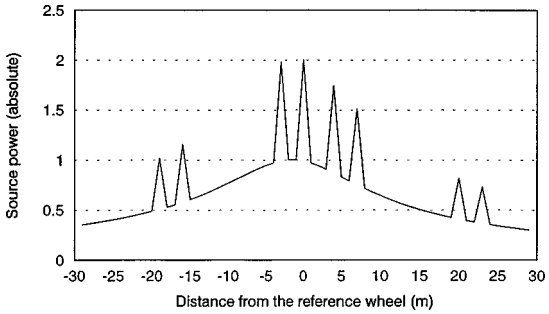


FIG. 5. Magnitude of source powers used to simulate the output from a train unit.

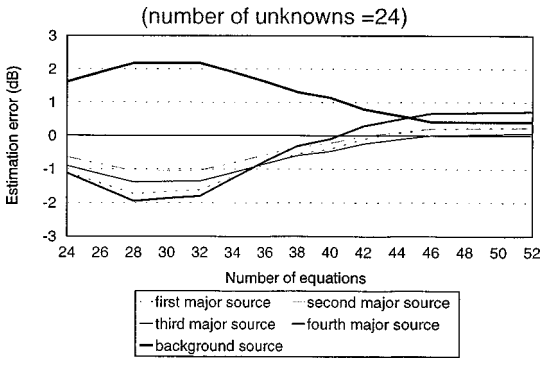


FIG. 6. Estimation errors against the number of equations.

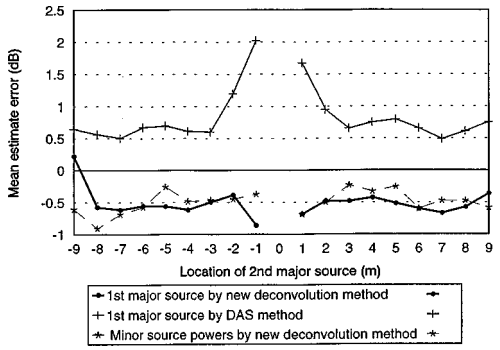


FIG. 7. Mean estimation error of the first major source (only two major sources were assumed).

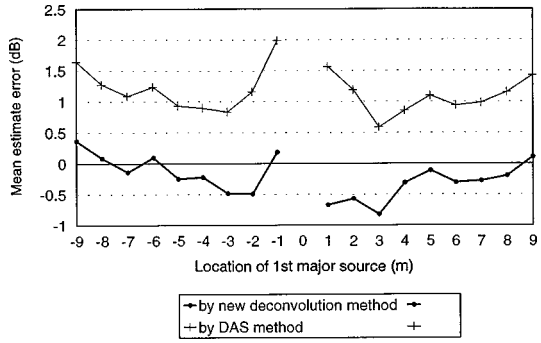


FIG. 8. Mean estimation error of the second major source (only two major sources were assumed).

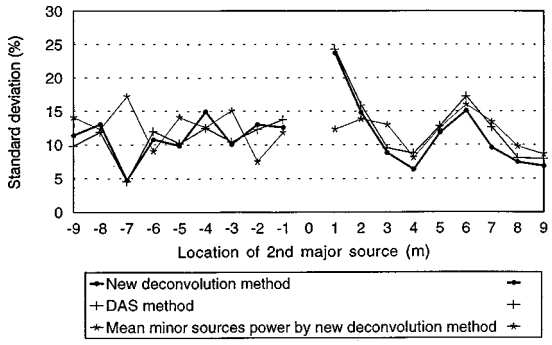


FIG. 9. Standard deviation of the estimates for the first major source (only two major sources were assumed).

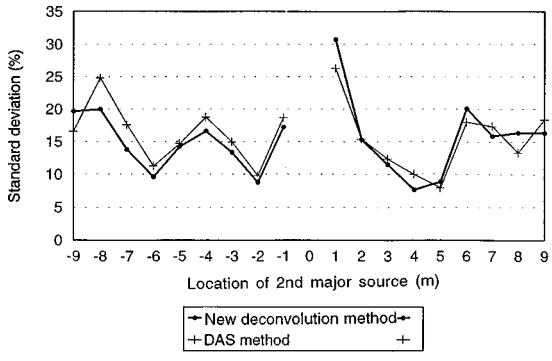


FIG. 10. Standard deviation of the estimates for the second major source (only two major sources were assumed).

TABLE I. Estimates obtained using the sweeping focus DAS method and the proposed deconvolution method. (Major sources at −8, −5, 0, 3 and 8 m.)

		Sources location (m)				
		−8	−5	0	+3	+8
Delay and sum	Mean estimate error (dB)	1.83	1.42	0.994	0.734	0.856
	Standard deviation (%)	22.6	23.3	18.6	13.2	16.35
Deconvolution	Mean estimate error (dB)	1.16	−0.128	−0.43	−0.62	1.035
	Standard deviation (%)	23.6	23.9	19.9	15.9	20.4

close to the dead ahead point can lead to large estimation errors. In order to obtain good power estimates using the deconvolution process, a careful selection of the system equations is vital. It is noted that only those equations representing the input–output relationship when the focus is steered close to the focus should be used.

The above discussion is related to the choice of the number of equations used in the deconvolution process. The number of discrete sources, Q (i.e., the number of variables in the process), should also be considered. It should be noted that the number of sources should not be greater than the number of equations. Moreover, all the major sources that would have significant contribution to the equations selected should be included. That means the optimal values of M and Q depend on the sources' configuration.

II. NUMERICAL SIMULATIONS

A number of simulations were performed to determine the right combination of M and Q. Figure 4 shows the detailed drawing of an electric multiple unit. A train will have eight electric multiple units. If the front wheel of the unit is taken as a reference point, there will be three other wheels located within 10 m from the reference point. Figure 5 shows the source locations and the modulated amplitudes used in these simulations. The major sources are located approximately at the wheel positions of a low speed electric train. The attenuation due to the distance effect is determined using the information available in Eq. (20). The number of variables was set to 24 and the four major sources within 10 m from the array axis were included. The power estimates of the major and minor sources were determined using the deconvolution process. Figure 6 shows the estimation errors for the different number of equations used. One can observe that the performance is the best when the number of equations is 44. However, the numbers of errors for different sources are slightly different.

To illustrate the advantages of using the proposed deconvolution method, simulations were performed on a moving line source. Both the sweeping focus DAS and the deconvolution methods were applied to this model. The model was idealized as 31 moving point sources lying on a straight line. They were located between −15 to +15 m with a separation between them of 1 m. They had a group velocity of 18 m/s. The source, which passed 0 m at $t=0$, was defined as the reference source. The signals of sources were generated by a 100-kHz random noise generator and filtered by a low pass filter of 5 kHz. The amplitude of each source was assigned according to its location at $t=0$ and was determined using Eq. (20). A random factor was multiplied to the amplitude to simulate the effect due to amplitude fluctuation. The random factor was set to 0.7–1.3. The amplitudes of the major sources were scaled up and the power ratio between the major to minor sources was set to 3 dB. It was assumed that the signals were received by an array of seven microphones with a separation of 0.2 m. The shortest distance between the sources and the array was set to 10 m. The sampling frequency of microphone was set to 10 kHz. The integration time was from −0.05 s to 0.05 s (corresponding to 1000 samples). The frequency of interest was up to 1 kHz. A total of ten realizations with independent sets of inputs were performed in each case.

Figures 7 and 8 show the mean power estimates of the ten realizations when there were only two major sources out of the 31 sources. One of the major sources was located at 0 m (which was named as the first major source) while the location of another major source (named as the second major source) was varied. When the two sources were situated at a close distance, their powers obtained using the sweeping focus DAS method were significantly overestimated. For the second major source located at a large angle with the array axis, the estimation error for the sweeping focus DAS method is large. This is because the mainlobe of the array

TABLE II. Estimates obtained using the sweeping focus DAS method and the proposed deconvolution method. (Major sources at −7, −4, 0, 5, and 8 m.)

		Sources location (m)				
		−7	−4	0	+5	+8
Delay and sum	Mean estimate error (dB)	2.13	1.84	1.06	1.29	1.66
	Standard deviation (%)	23.9	16.8	16.3	16.4	12.6
Deconvolution	Mean estimate error (dB)	1.69	0.43	−0.40	0.29	1.59
	Standard deviation (%)	29.3	18.2	14.2	14.3	24.3

TABLE III. Estimates obtained using the sweeping focus DAS method and the proposed deconvolution method. (Major sources at −8, −2, 0, 6, and 9 m.)

		Sources location (m)				
		−8	−2	0	+6	+9
Delay and sum	Mean estimate error (dB)	1.17	1.54	1.62	1.43	1.51
	Standard deviation (%)	12.1	22.7	23.0	22.0	15.7
Deconvolution	Mean estimate error (dB)	1.12	−0.59	−0.17	0.78	1.27
	Standard deviation (%)	18.2	20.4	19.1	26.8	28.1

pattern in the sweeping focus DAS method becomes wider when the array is steered to a large angle with the array axis, and thus inevitably would include more sources in the lobe. In general, the powers obtained using the deconvolution method are much more accurate. The minor source power level means were only underestimated by about 0.5 dB for various major source locations. Figures 9 and 10 show the standard deviation of the estimates of the above 10 realizations, which were normalized with respect to the mean level. One can observe that the normalized standard deviations are about the same for both methods.

Some other simulations with five major sound sources at various locations were performed to test the applicability of the proposed deconvolution method on power estimation. Tables I–III show the mean results of the estimates obtained using the sweeping focus DAS and the proposed deconvolution methods. One can observe that the mean estimate errors are small for those sources located close to the dead ahead point. The power means obtained using the deconvolution method are more accurate than those obtained using the sweeping focus DAS method only. Both methods give similar magnitude for standard deviations.

III. TRAIN NOISE MEASUREMENT USING A MICROPHONE ARRAY

The microphone array consisted of 15 Cirus microphones with windscreens and the separation between microphones was set to 0.2 m. The shortest distance between the array and the track was 7.8 m at which the effect due to the boundary layer induced by the movement of trains was negligible. The array was mounted horizontally at 1.2 m above the ground and has an alias frequency of 1700 Hz. Due to the sound refraction caused by the boundary layer, the apparent sound sources located by the array would be shifted. Barsikow *et al.*[5] showed the shift to be less than 5 cm at a train speed of 60 km/h. Therefore, the shift was neglected in the study.

The microphone signals, after being amplified, were recorded by a SONY 16-channel digital audio tape (DAT) recorder with a dynamic range of 80 dB. A calibration signal from a B&K 4231 acoustic calibrator was recorded before and after each noise measurement. A pair of full-beam photoelectric sensors were mounted on both sides of rail tracks. They would produce a pulse signal when a wheel cut through the light beam of sensor and the signal was also recorded by the DAT recorder. The measured sound pressure signals could then be correlated with the movement of trains. The train speed was determined by using the separation of wheels and the duration of a pulse. In total 15 values of train speed were calculated for one train. The train speed was about 60 km/h.

Figure 11 shows a typical broadband steering response of a moving train using the DAS method with sweeping focus. The tracking period of focus was 0.125 s (1500 samples), corresponding to ±1.1 m along the train. The sweeping procedure was repeated at every 0.3 m along the train. The time history obtained by the center microphone is also shown in Fig. 11. A comparison of the results obtained using the DAS method and a single microphone shows there are about 10-dB difference. This is because, as stated in Sec. I A, the DAS method can act as a spatial filter and most of the sound from directions other than the focus can be excluded. In contrast, a single microphone can integrate the sound from all directions. Most peaks were located at the wheel positions. The difference between the peak levels (most of them at the wheel locations) and the nonpeak levels at the middle of the train unit is about 5–10 dB for the DAS

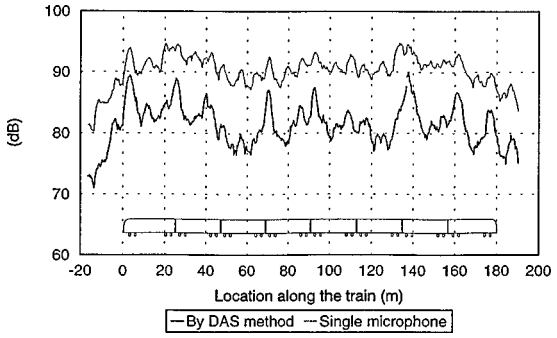


FIG. 11. Steering response of DAS and the time history of a single microphone.

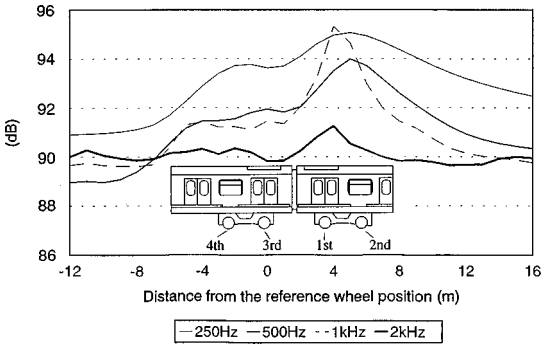


FIG. 12. Mean steering response zooming at the wheel locations.

TABLE IV. Comparison of mean power estimates and standard deviation of estimates at different wheel locations and at the middle positions of the EMU (250-Hz octave band).

		Sources location				
		1st wheel	2nd wheel	3rd wheel	4th wheel	Middle of an EMU
Power estimates (dB)	Delay and sum	95.1	95.0	96.3	96.3	92.7
	Deconvolution	91.7	91.9	92.9	93.6	88.7
Standard deviation (%)	Delay and sum	37.6	29.3	35.9	39.5	19.7
	Deconvolution	75.2	55.8	85.5	68.9	26.1

TABLE V. Comparison of mean power estimates and standard deviation of estimates at different wheel locations and at the middle positions of the EMU (500-Hz octave band).

		Sources location				
		1st wheel	2nd wheel	3rd wheel	4th wheel	Middle of an EMU
Power estimates (dB)	Delay and sum	92.9	93.3	94.9	95.1	90.9
	Deconvolution	91.1	91.0	94.1	93.6	88.7
Standard deviation (%)	Delay and sum	34.8	30.1	30.5	28.2	19.9
	Deconvolution	73.1	47.5	49.5	55.6	23.8

TABLE VI. Comparison of mean power estimates and standard deviation of estimates at different wheel locations and at the middle positions of the EMU (1000-Hz octave band).

		Sources location				
		1st wheel	2nd wheel	3rd wheel	4th wheel	Middle of an EMU
Power estimates (dB)	Delay and sum	92.6	92.9	96.7	94.3	92.1
	Deconvolution	90.4	91.0	93.0	90.7	89.5
Standard deviation (%)	Delay and sum	29.6	31.9	38.1	38.4	35.7
	Deconvolution	58.8	42.4	81.7	50.3	26.8

TABLE VII. Comparison of mean power estimates and standard deviation of estimates at different wheel locations and at the middle positions of the EMU (2000-Hz octave band).

		Sources location				
		1st wheel	2nd wheel	3rd wheel	4th wheel	Middle of an EMU
Power estimates (dB)	Delay and sum	91.5	91.0	92.5	91.5	90.7
	Deconvolution	88.9	89.5	92.7	91.2	90.4
Standard deviation (%)	Delay and sum	20.3	26.5	33.2	25.0	24.2
	Deconvolution	31.1	35.2	39.6	41.2	25.2

method while this difference is only about 3–5 dB for the single microphone.

Figure 12 shows the mean steering response zooming at the wheel positions for various octave bands. As the array has an alias frequency of 1700 Hz, the results obtained using the DAS for the highest frequency octave band would contain unwanted components of sound measured with the alias mainlobes. Nevertheless, this figure shows that peaks are always located at the wheel positions, especially in the low- and mid-frequency ranges.

It has been shown in Sec. II that for moving trains, the number of equations and variables in the deconvolution process could be set to 24 and 44, respectively, in order to produce accurate power estimates. Tables IV–VII show the results of power estimates obtained using the sweeping focus DAS and the proposed deconvolution method. In general, the estimates obtained using the sweeping focus DAS method are larger than those obtained using the deconvolution method as anticipated. One can explain this phenomenon by using Rayleigh criterion, $\sigma = \lambda/L$, where σ is the resolution in radians, λ is the acoustic wavelength, and L is the length of the array for the DAS method. For the array with microphones separated by an interval of 0.2 m and the frequency of interest at 1000 Hz, the array would have a resolution of approximately 0.28 rad or when projected onto the line of sources about 2.8 m ($>$2.5 m for separation between two adjacent wheels). The estimates obtained using the proposed deconvolution method would be much more accurate because the deconvolution method can recover the source powers by using train source characteristics. The difference between the power estimates of major sound sources obtained using the two methods decreases when the frequency of interest increases because the mainlobe of the sweeping focus DAS array pattern is narrower in the high-frequency range.

In Sec. II, numerical results showed that the standard deviations of power estimates obtained using the deconvolution method were slightly larger (less than 5%) for various cases. However, the measured data shown in Tables IV–VII indicated that the standard deviations of power estimates obtained using the deconvolution method were, in general, much larger than those obtained using the sweeping focus DAS method. The main reason is that the sweeping focus DAS method has a relatively wide mainlobe, which will enhance the averaging effect for a number of trains and subsequently will produce small standard deviations in the low-frequency range. Moreover, the large difference of standard deviations obtained using the deconvolution method could reveal the variation on the condition of wheels which might emit high sound levels in the low-frequency range.

IV. CONCLUSION

This paper presents a deconvolution approach that can estimate the sound powers at various parts on a moving line source. The traditional sweeping focus delay and sum (DAS) method and the proposed deconvolution method were applied to a series of numerical simulations. Simulation results showed that the proposed deconvolution method would produce smaller mean estimation errors and could be used to deal with cases when the separation of major sources was small.

The sweeping focus DAS method and the deconvolution method were applied to the recorded noise data emitted from eight electric trains. An array of 15 microphones with a pair of full-beam photoelectric sensors was built to record the noise data as well as the location of moving trains. The sweeping focus DAS method was used to estimate the directions of sound sources and then the sound powers were estimated by both the sweeping focus DAS method and the deconvolution method. The sweeping focus DAS method always overestimated the sound powers on moving trains while the deconvolution method would give better estimates of the sound powers on moving trains. Moreover, the large difference of standard deviations obtained using the deconvolution method reveals the variation on the condition of wheels that emit high sound levels in the low-frequency range.

ACKNOWLEDGMENT

The authors would like to thank C. W. Kwok for his assistance during the preparation and execution of the measurements.

[1] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing* (Prentice–Hall, Englewood Cliffs, NJ, 1993).
[2] J. Billingsley and R. Kinns, ''The acoustic telescope,'' J. Sound Vib. **48**, 485–510 (1976).
[3] W. F. King III, ''On the role of aerodynamically generated sound in determining wayside noise levels from high speed trains,'' J. Sound Vib. **54**, 361–378 (1977).
[4] W. F. King III and D. Bechert, ''On the sources of wayside noise generated by high speed trains,'' J. Sound Vib. **66**, 311–332 (1979).
[5] B. Barsikow, W. F. King, and E. Pfizenmaier, ''Wheel/rail noise generated by a high speed train investigated with a line array of microphones,'' J. Sound Vib. **118**, 99–122 (1987).
[6] B. Barsikow, ''Experiences with various configurations of microphone arrays used to locate sound sources on railway trains operated by the DB AG,'' J. Sound Vib. **193**, 283–293 (1995).
[7] C. C. Ko and H. Liu, ''Estimation of powers of uncorrelated sources in linear arrays,'' Signal Process. **46**, 245–248 (1995).
[8] R. Roy and T. Kailath, ''ESPRIT-estimation of signal parameters via rotational invarience techniques,'' IEEE Trans. Acoust., Speech, Signal Process. **ASSP-37**, 984–995 (1989).
[9] S. S. M. Yung and W. M. To, ''Localization of noise sources of a low speed electric train,'' Proc. Asia-Pacific Vibration Conference '95 **2**, 441–446 (1995).
[10] J. Lardies, ''Power estimation of multiple sources by an acoustical array,'' J. Sound Vib. **145**, 309–320 (1991).
[11] J. J. Fuchs and H. Chuberre, ''A deconvolution approach to source localization,'' IEEE Trans. Signal Process. **42**, 1462–1470 (1994).
[12] H. Chuberre and J. J. Fuchs, ''A deconvolution approach to moving source localization,'' Proc. IEEE Conference on Acoustics, Speech and Signal Processing **2**, 349–352 (1994).
[13] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).

A multiple-frequency method for potentially improving the accuracy and precision of *in situ* target strength measurements

David A. Demer
Southwest Fisheries Science Center, P.O. Box 271, La Jolla, California 92038

Michael A. Soule
Sea Fisheries Research Institute, Private Bag X2, 8012 Rogge Bay, Cape Town, South Africa

Roger P. Hewitt
Southwest Fisheries Science Center, P.O. Box 271, La Jolla, California 92038

(Received 11 May 1998; revised 13 November 1998; accepted 4 January 1999)

The effectiveness of a split-beam echosounder system to reject echoes from unresolvable scatterers, thereby improving the measurements of *in situ* target strengths (TS) of individuals, is dramatically enhanced by combining synchronized signals from two or more adjacent split-beam transducers of different frequencies. The accuracy and precision of the method was determined through simulations and controlled test tank experiments using multiple standard spheres and 38- and 120-kHz split-beam echosounders. By utilizing the angular positional information from one of the split-beam transducers, additional corresponding TS measurements were shown to be obtainable from a juxtaposed single-beam transducer. Both methods were utilized to extract *in situ* TS measurements of Antarctic scatterers simultaneously at 38, 120, and 200 kHz. The ultimate efficiency of the multiple-frequency method is shown to be limited by phase measurement precision, which in turn is limited by the scattering complexity of targets, the signal-to-noise ratio, and the receiver bandwidth. Imprecise phase measurements also result in significant beam-compensation uncertainty in split-beam measurements. Differences in multi-frequency TS measurements provided information about the identity of constituents in a mixed species assemblage. The taxa delineation method has potential, but is limited by compounding measurement uncertainties at the individual frequencies and sparse spectral sampling. © *1999 Acoustical Society of America.*
[S0001-4966(99)01204-7]

PACS numbers: 43.60.Qv, 43.30.Vh, 43.30.Xm [JCB]

INTRODUCTION

A distribution of backscattering cross-sectional areas of individual scatterers ($P\{\sigma_{bs}\}$) is a critical factor in the estimation of animal density (numbers/km^2) from an acoustical survey (Bodholt, 1990; Hewitt and Demer, 1993). Measurements of σ_{bs} (m^2/animal), frequently expressed as target strength [$\mathrm{TS}=10\log(\sigma_{bs})$], can be made either in their natural state and environment (*in situ*) or through controlled experimentation (Ehrenberg, 1989). However, the TS of fish and zooplankton are highly dependent upon the variation of sound scattering with animal size, shape, orientation, and acoustic impedance (Chu *et al.*, 1992; Stanton *et al.*, 1994; Demer and Martin, 1995). Thus, the applicability of TS measurements made on constrained, sedated, or dead animals may be suspect. On the other hand, efforts to make direct *in situ* TS measurements have been hindered by equipment and physical limitations (Ehrenberg, 1979; Foote, 1991). This study aims to improve the methods for measuring TS *in situ*, concurrent with echo-integration surveys, by requiring simultaneously detected echoes to pass multiple-target rejection algorithms at two or more frequencies. Additionally, the simultaneous measurements of TS at multiple frequencies are used to empirically characterize the frequency-dependent scattering or ''acoustic signatures'' of some species. In some cases, the latter may permit acoustical identification of taxa and the apportionment of total echo energy to different species.

The echosounder used in this study is the three-frequency Simrad EK500 (Bodholt *et al.*, 1988). A series of simulations (Sec. I) and test tank experiments (Sec. II) were conducted to (1) characterize the precision of TS measurements made with the EK500 at 38 and 120 kHz; (2) evaluate the adequacy of the EK500 single-frequency algorithm for rejecting multiple targets in some situations; and (3) demonstrate that the individual scatterers are more accurately delineated by matching the outputs of the EK500 algorithm at two or more frequencies. Field experiments (Sec. III) were then conducted to demonstrate the efficacy of the new method for accurately and simultaneously measuring *in situ* TS measurements of individual scatterers at three frequencies (38, 120, and 200 kHz). With adjacently mounted transducers, the target-position information from the split-beam systems was used to derive beam-compensated TS values from the 200-kHz single-beam system (Demer, 1994). Finally, differences in the simultaneous TS measurements were used to relate the acoustic measurements to the species composition of net samples.

In the following paragraphs, several terms or concepts (animal density, split-beam measurements, noise effects, single target detections, and taxa delineation) are described as they pertain to fisheries acoustics theory and instrumentation.

A. Animal density estimation

In acoustical surveys of aquatic organisms, the areal density of scatterer type ''x'' (ρ_x) can be estimated from the integrated volume backscattering strength (S_{A_x}), or the total backscattering cross-sectional area per unit of sea-surface area (m^2/km^2) from type-x animals, divided by the frequency-weighted backscattering cross-sectional area from an individual type-x animal ($P\{\sigma_{\mathrm{bs}_x}\}\cdot\sigma_{\mathrm{bs}_x}$). In practice, a discrete sum form of the probability density function (PDF) is used:

$$\rho_x = S_{A_x} \Big/ \sum_{i_x=1}^{n_2} f_{x_i}\sigma_{\mathrm{bs}_{x_i}}, \tag{1}$$

where f_{x_i} is the relative frequency of type-x animals with $\sigma_{\mathrm{bs}_{x_i}}$, such that $\Sigma_{i_x=1}^{n_x} f_{x_i}=1$, where i_x refers to the ith class of σ_{bs_x} and n_x is the number of σ_{bs}-classes of type-x animals.

B. Split-beam measurements

During the survey, *in situ* TS measurements of individual scatterers are possible using a split-beam echosounder to locate echoes in three-dimensional space and a method for excluding unresolvable targets. A sound pulse is transmitted simultaneously from four quadrants of a split-beam transducer and received by each quadrant individually. The electrical phase between the signals received from two halves of the transducer (ϕ_e) is related to the angle between the beam-axis and the target (θ):

$$\phi_e = kd_{\mathrm{eff}}\sin\theta, \tag{2}$$

where k is the acoustic wave number and d_{eff} is the effective separation between two transducer halves (a function of element shading). In the alongship plane, $\theta=\alpha$ and $d_{\mathrm{eff}}=d_\alpha$ and in the athwartship plane, $\theta=\beta$ and $d_{\mathrm{eff}}=d_\beta$. The angular resolution of the system (ζ) is determined by differentiating ϕ_e with respect to θ:

$$\zeta = \frac{\Delta\phi_e}{kd_{\mathrm{eff}}\cos\theta}, \tag{3}$$

and evaluating with the phase resolution of the echosounder (in the case of the EK500, $\Delta\phi_e=180$ electrical degrees per 64 phase steps), a small angle approximation ($\cos\theta\cong 1$), and the transducer angle sensitivity ($\Lambda=kd_{\mathrm{eff}}$):

$$\zeta \cong \left(2.8125\left(\frac{\text{electrical degree}}{\text{phase step}}\right)\right) \Big/ \left(\Lambda\left(\frac{\text{electrical degree}}{\text{spatial degree}}\right)\right). \tag{4}$$

The radial range to the target ($r=ct_{2\text{-way}}/2$) is estimated from the sound speed (c) and half of the two-way propagation delay ($t_{2\text{-way}}$). Uncertainty in the determination of $t_{2\text{-way}}$ is affected by the echo pulse rise time and a small delay in the receiving electronics (MacLennan, 1987). Therefore, $t_{2\text{-way}}$ is most accurately estimated by subtracting the system-dependent receiver delay (t_{del}) from the propagation delay (t_h) measured between the start of the transmit pulse and the point on the leading edge of the echo at which the amplitude has risen to half the peak value (MacLennan and Simmonds, 1992; Fernandes and Simmonds, 1996).

Thus, the split-beam system allows an individual animal to be located in three-dimensional space (α, β, and r), provided that it is separated from other scatterers by a radial distance greater than the range resolution or half the product of c and the pulse duration (τ) (MacLennan and Simmonds, 1992). Of course, the echo amplitude must also exceed the detection threshold. In this case, the backscattering cross-sectional area of an individual scatterer can be estimated from the ratio of the received power (p_r) to the transmit power (p_t), with compensations for spherical spreading, absorption, system gain, and the transducer beam pattern (Kerr, 1988; Simrad, 1996):

$$\sigma_{\mathrm{bs}} = \frac{p_r 16\pi^2 r^4 10^{2\gamma r}}{p_t g_0^2 r_0^2 \lambda^2}\cdot b(\alpha,\beta)^2, \tag{5}$$

where γ is the absorption coefficient (dB/m), g_0 is the system gain on the beam axis, r_0 is the reference distance (1 m), and $b(\alpha,\beta)$ is the one-way beam pattern compensation evaluated in the direction of the target [$b(\alpha,\beta)\geqslant 0$]. In logarithmic form,

$$\begin{aligned}\mathrm{TS} = {} & P_r + 20\log(4\pi) + 40\log(r) + 2\gamma r - P_t - 2G_0 \\ & - 20\log(\lambda) + 2B(\alpha,\beta),\end{aligned} \tag{6}$$

where the decibel forms of p_t, p_r, g_0, and $b(\alpha,\beta)$ are P_t, P_r, G_0, and $B(\alpha,\beta)$, respectively (dB *re*: 1 W). The first seven terms of (6) are collectively termed the uncompensated target strength (TS_U):

$$\mathrm{TS} = \mathrm{TS}_U + 2B(\alpha,\beta). \tag{7}$$

Knowing the off-axis angles and the shape of the beam, compensation can be applied to normalize all TS measurements to the calibrated beam axis.

C. Noise effects

As with all acoustical measurements, background noise (e.g., thermal, wind, bubble, engine, tank boundary reverberation, etc.) can affect both the accuracy and precision of the phase measurements and thus the estimates of off-axis angles. The total sound pressure (p) is the sum of the complex signal (p_s) and the complex noise (p_n), with time-varying amplitudes [$P_s(t)$ and $P_n(t)$, respectively]:

$$p_s = P_s(t)e^{j\omega t} = P_s(t)(\cos(2\pi ft) + j\sin(2\pi ft)), \tag{8}$$

$$p_n = P_n(t)e^{j\omega t} = P_n(t)(\cos(2\pi ft) + j\sin(2\pi ft)), \tag{9}$$

$$\begin{aligned} p &= (P_s(t) + P_n(t))e^{j\omega t} \\ &= (P_s(t) + P_n(t))(\cos(2\pi ft) + j\sin(2\pi ft)).\end{aligned} \tag{10}$$

The phase estimate from the qth quadrant ($\hat{\phi}_{e_q}$) is determined by the arctangent of the imaginary part of the total received pressure divided by the real part (Senturia and Wedlock, 1975). Assuming the signal has zero phase (arbitrary reference) and a constant amplitude which is much greater than the noise level [$P_s \gg P_n(t)$], a small angle approximation can be employed [$\tan^{-1}(\phi_{e_q})\cong\phi_{e_q}$], and the phase estimate can be approximated by

TABLE I. Minimum SNR (dB) to estimate off-axis angles ($\hat{\theta}=\alpha$ or β) which are precise (95% confidence interval) to within one to five quantization steps ($\Delta\phi_e$). The requisite SNR decreases as the number of phase samples (Z) in the average phase difference estimate ($\overline{\Delta\hat{\phi}}$) increases. The basic sampling rates of the EK500 are 7.5, 25, and 37.5 kHz at 38, 120, and 200 kHz, respectively. Therefore, for pulse durations of 1.0 ms at 38 kHz and 0.3 ms at 120 kHz, approximately seven samples are averaged and a SNR of 24 dB is required for maximum precision in the phase measurements.

Minimum signal-to-noise ratio (dB)				SNR for phase estimate ($\Delta\hat{\phi}_e$)	SNR for average estimated phase ($\overline{\Delta\hat{\phi}_e}$)								
Phase (θ)				Samples in average (Z)									
Radians ($s\Delta\phi_e$)	Degrees (ES3812)	Degrees (ES120-7)	Steps (s)	1	2	3	4	5	6	7	8	9	10
0.049	0.23	0.13	1	32	29	27	26	25	24	24	23	23	22
0.098	0.45	0.27	2	26	23	21	20	19	18	18	17	17	16
0.147	0.68	0.40	3	23	20	18	17	16	15	14	14	13	13
0.196	0.90	0.54	4	20	17	15	14	13	12	12	11	11	10
0.245	1.13	0.67	5	18	15	13	12	11	10	10	9	9	8

$$\hat{\phi}_{e_q} \cong \frac{P_n(t)\sin(2\pi ft)}{P_s\cos(2\pi ft)}. \tag{11}$$

Assuming the noise is Gaussian with root mean square (rms) intensity σ_n^2 (Greenlaw and Johnson, 1983), and recalling that $\mathrm{var}[Cx]=C^2\,\mathrm{var}[x]$, the variance of the phase estimate is

$$\mathrm{var}[\hat{\phi}_{e_q}]=\mathrm{var}\left[\frac{P_n\sim N(\overline{x},\sigma_n^2)}{P_s}\right]=\frac{1}{P_s^2}\mathrm{var}[P_n(t)]=\frac{\sigma_n^2}{P_s^2}, \tag{12}$$

and the signal-to-noise power ratio (snr) is

$$\mathrm{snr}=\frac{P_s^2}{2\sigma_n^2}. \tag{13}$$

The variance of the single-channel phase estimate is determined by solving (13) for σ_n, substituting into (12):

$$\mathrm{var}[\hat{\phi}_{e_q}]=\frac{1}{2\ \mathrm{snr}}. \tag{14}$$

To measure phase in the EK500, the four receiver channels are amplitude limited to allow comparison of the delay between zero crossings for phase determination. The measured delays are interpreted in firmware as phase differences between half-beams. Depending upon the clock-rate and the algorithm, more than one estimate of single-channel phase may be averaged for each estimate of the phase difference between half-beams. Moreover, relative phases may be more precisely determined by cross correlation (Medwin and Clay, 1998). Conservatively, however, assume that each phase-difference estimate ($\Delta\hat{\phi}_e$) is determined from two unaveraged estimates of single-channel phase ($\hat{\phi}_{e_1}-\hat{\phi}_{e_2}$); therefore, $\mathrm{var}(\Delta\hat{\phi}_e)=2\,\mathrm{var}(\hat{\phi}_{e_q})$. The actual phase estimates are averages of Z independent phase difference estimates made within the -6-dB points of the echo envelope ($\overline{\Delta\hat{\phi}_e}=(\Sigma_Z(\Delta\hat{\phi}_e)_Z)/Z$) and $\mathrm{var}[\overline{\Delta\hat{\phi}_e}]=\mathrm{var}[\Delta\hat{\phi}_e]/Z=(2/Z)\cdot\mathrm{var}[\hat{\phi}_{e_q}]$ (Rice, 1988). Thus, for maximum precision of the mean phase difference between half-beams, the phase quantization of the echosounder ($\Delta\phi_e=0.0491$ radians/phase step) must be greater than $2\cdot sd[\hat{\phi}_{e_q}]$ (95% confidence interval):

$$\mathrm{SNR}>10\log\left(\frac{4}{Z(s\Delta\phi_e)^2}\right), \tag{15}$$

where $\mathrm{SNR}=10\log(\mathrm{snr})$ in decibels (dB) and s is the integer number of phase steps. For maximum precision, s is 1. If reduced precision is acceptable, (15) can be evaluated with increased numbers of phase steps (Table I).

D. Single-target detection

In order to reject returns from unresolvable targets, a multi-tiered algorithm can be employed (Bodholt, 1991; Bodholt and Solli, 1992). As implemented in the EK500 Echosounder (Simrad, 1996), filters are imposed on the minimum TS, the minimum and maximum duration of the echo envelope (echo width at -6-dB points normalized by τ), maximum $B(\alpha,\beta)$, and the maximum sample-to-sample deviation of the phase measurements (Table II). The constraint on the minimum normalized echo length is intended to reject noise and destructively interfering echoes from multiple targets with different bearings but similar ranges. The maximum normalized echo length criteria is intended to eliminate overlapping echoes from multiple targets at similar bearings, but different yet unresolvable ranges. The maximum gain compensation is intended to reduce biases due to thresholding of smaller targets at the beam periphery (reduced sensitivity), and a decrease in the accuracy of the estimated beam compensation function $[\hat{B}(\alpha,\beta)]$ at large off-axis angles. The maximum phase deviation is employed to reject incoherent echoes from unresolvable target multiples. In the EK500 firmware versions 5.0 and later, an improved phase deviation classifier (P_{dev}) is employed which is simply the standard deviation of the phase samples within the echo length (Soule *et al.*, 1997).

Unfortunately, the multiple-target rejection algorithm is sometimes imperfect at high scatterer densities (Hewitt and Demer, 1991). This is especially true when constructive-interference results from multiple scatterers residing in the

TABLE II. Multiple-target rejection criteria used in both the tank and field experiments. Unresolvable targets are filtered, to some extent, by imposing (1) a minimum target strength (Min. TS); (2) minimum and maximum duration of the echo envelope, normalized by the pulse duration (τ); (3) maximum one-way beam compensation (Max. B); and (4) a maximum sample-to-sample deviation of the phase measurements (P_{dev}).

Minimum target strength (TS)	-90 dB
Minimum echo length	0.8τ (s)
Maximum echo length	1.5τ (s)
Maximum beam compensation [$B(\alpha,\beta)$]	4 dB[a]
Maximum phase deviation (P_{dev})	4 steps[a]

[a]Not applicable for single-beam system.

same resolution volume (Soule *et al.*, 1995, 1996); their simulations and tank measurements showed that overlapping echoes from targets separated by half-multiples of a wavelength were preferentially accepted. However, because the interference is dependent upon the relationships between acoustic wavelength ($\lambda = c/f$) and scatterer spacing, the effects vary with frequency (Foote, 1996). In other words, for the range separation between two targets (Δr) to be equal to an integer multiple of $\lambda/2$ at two different discrete frequencies simultaneously, one wavelength (λ_1) must be M/N times the other wavelength (λ_2), where M and N are both integers. This occurs very infrequently. For example, at the precise EK500 frequencies of 37.879 and 119.048 kHz, a sound speed (c) of 1500 m/s, pulse lengths of 1.0 ms, and transducer beamwidths of 7.0°, this exact situation does not occur within the resolution volume, bounded by the pulse width and the transducer -3-dB points, to a maximum range of 250 m. Compromising this ideal situation are the finite bandwidths of the pulsed transmissions ($\cong 1$ kHz) which may still allow interference to occur when target spacing is slightly more or less than a half-wavelength of the center frequency. Nevertheless, it follows that the accuracy and precision of *in situ* TS measurements can likely be improved by requiring simultaneously detected echoes to pass multiple target rejection algorithms at two or more frequencies.

E. Taxa delineation

In addition to reducing measurement uncertainty, simultaneous measurements of *in situ* TS at more than one frequency may provide enough information to identify scatterers. Holliday (1977) showed that biophysical information could be extracted from multi-frequency acoustic backscatter. Greenlaw (1979) then proposed, and Madureira *et al.* (1993) and Brierly *et al.* (submitted) demonstrated, that differences in volume backscattering strength measured at multiple frequencies could be used to classify scatterer sizes. Greenlaw's method was also adapted to multi-frequency *in situ* TS measurements, and coupled with distributions of theoretical expectations (e.g., $P\{\sigma_{120} - \sigma_{38\,\text{kHz}}\}$), to estimate taxa of individual scatterers (Demer, 1994).

I. SIMULATIONS

The potential utility of the multiple-frequency method for rejecting echoes from multiple unresolvable targets was first tested with a simulation similar to that conducted in Soule *et al.* (1996). Two unresolvable targets were randomly and independently positioned within a three-dimensional pulse resolution volume (at 120 kHz), and assigned a point-location in a spherical coordinate system ($10.0 \leqslant r \leqslant 10.75$ m; 0 degrees$\leqslant$polar angle$\leqslant$3 degrees; and 0 degrees$\leqslant$azimuth angle $\leqslant$360 degrees). Transmitted pulses were simulated at both 38 and 120 kHz by multiplying a Gaussian envelope and a continuous sine wave at each frequency. The band-limited pulses had amplitudes of unity and were approximately 1.0 ms between -6-dB points. The complex echo waveforms were computed as received by each quadrant of collocated 38 and 120 kHz split-beam transducers. The transmission losses were neglected and the sampling rates were chosen to be equivalent to the EK500 (7.5 and 25 kHz, respectively). Both transducers were modeled as point-receivers with ideal beams. Ignoring shading, quadrant separations of 138.5 and 55 mm resulted in angle sensitivities of 21.9 and 23.8 at 38 and 120 kHz, respectively.

At each frequency, the reflected pressure waves received at each of the four transducer elements ($q = 1$, 2, 3, or 4) from each of the two randomly located targets ($l = 0$, 1, or 2) is given by Soule *et al.* (1996):

$$p_{q1}(t) = \sigma_{\text{bs}_1}^{1/2} \cdot e^{(-[t-((r_{q1}+r_1)/c)]^2/2T_p^2)} \cdot e^{j(\omega t-(r_{q1}+r_1)k)}, \qquad (16)$$

where σ_{bs_1} is the backscattering cross-sectional area of the lth target, r_{q1} is the range from element q to target 1, r_1 is the range from the origin to target l, $c = 1500$ m/s, j is the imaginary number ($\sqrt{-1}$), T_p is the pulse width parameter (0.4 ms) used to set the duration of the Gaussian envelope, k is the wave number ($2\pi c/f$), and the radian frequency ($\omega = 2\pi f$) is assumed to be a constant for narrow bandwidth operation. Individual backscattering cross-sectional areas were generated from a log-normal distribution [$N(\overline{x},\sigma^2)$], with mean $\overline{x} = 0$ dB and standard deviation $\sigma = 3$ dB.

At each frequency, echoes were rejected as being multiples if the standard deviation in phase samples (P_{dev}), recorded within 6 dB of the peak, exceeded a preset limit in either the athwartship or alongship directions (Bodholt, 1990). Consistent with the minimum echo length filter, echoes were also rejected if the trough in a multi-peak echo was more than 6 dB below the peak. Runs of 1000 pings were repeated at P_{dev} limits corresponding to off-axis angles of 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, and 1.0 degrees.

The echoes which passed these filters at a single-frequency were then subjected to two-frequency spatial matching criteria; echoes were rejected if (1) the range estimates at the two frequencies were not matched to within 0.1 m; and (2) the absolute difference between the respective estimates of off-axis angles (α and β) were not matched to within a specified angle discrepancy.

In some cases, unresolvable targets are falsely identified as individual scatterers by a split-beam system utilizing a single-frequency multiple-target rejection algorithm (see Sec. D of the Introduction). The deleterious nature of this problem has been clearly demonstrated [Fig. 1(a)]. Simulations of the one-frequency multiple-target rejection algorithm show a dramatic increase in the number of unresolv-

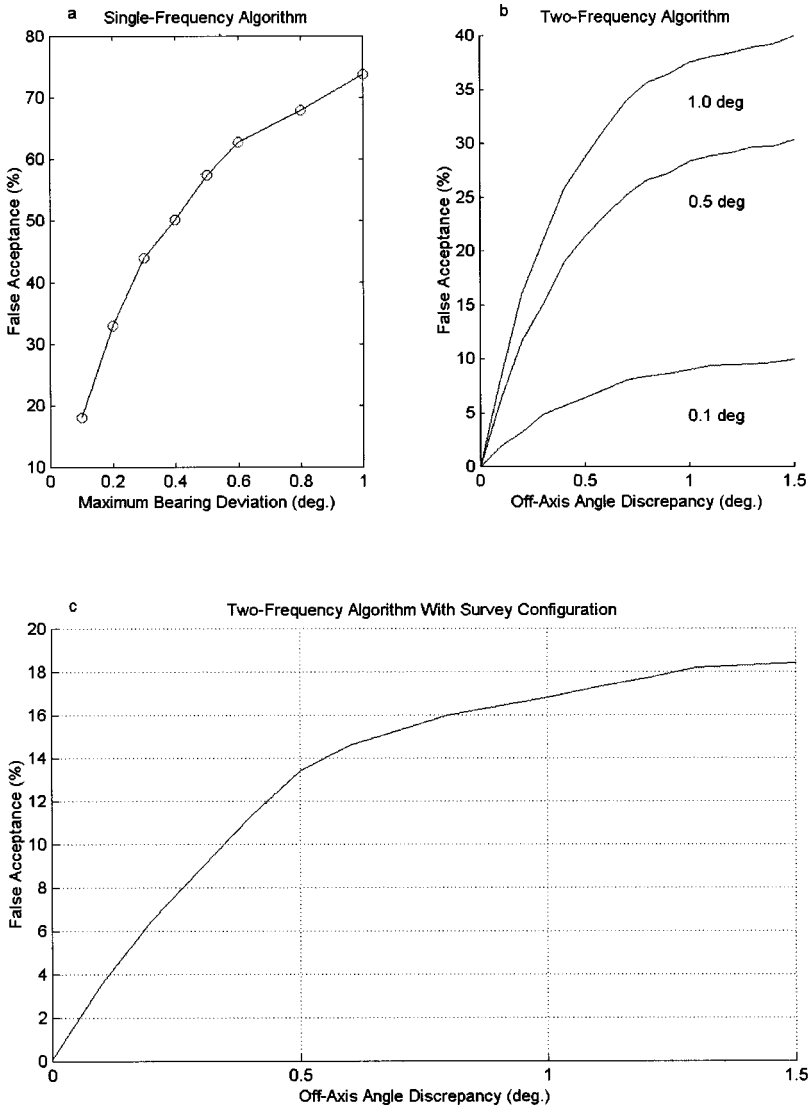


FIG. 1. Simulated false detection rates (acceptance of unresolvable scatterers as individuals) using the one- and two-frequency multiple-target rejection algorithms. With a single frequency (38 or 120 kHz), the percent of false detections rises sharply with an increase in maximum bearing deviation (electrical phase deviation converted to off-axis angles). At the default phase deviation of four steps (equivalent to a bearing deviation $A_{\mathrm{dev}} \cong 0.52$ degrees for the Simrad ES120-7 120-kHz transducer), the simulation predicts a false acceptance of about 57.9%. Using the two-frequency method (b), the false acceptance rate is dramatically reduced, even broad limits on the off-axis angle discrepancy, to less than 40%, 30%, and 10% with A_{dev} at 1.0, 0.5, and 0.1 degrees, respectively. Under the survey equipment configuration (notably, pulse lengths of 1.0 and 0.3 ms at 38 and 120 kHz, respectively), using default EK500 filter parameters (notably $A_{\mathrm{dev}} \cong 0.52$ degrees, 75.4% of the multiple targets were falsely accepted at 38 kHz, 25.3% at 120 kHz. However, using the two-frequency method (range and angle-discrepancy criteria of 0.1 to 1.5 degrees, respectively), the simulation predicts false acceptances of only 3.6 % to 18.4 % using (c). Note the scale changes.

able targets accepted with an increase in maximum bearing deviation criteria [electrical phase deviation (P_{dev}) converted to off-axis angle deviation (A_{dev})]. Using the EK500 default phase deviation setting (Table II, $P_{\mathrm{dev}}=4$ steps or $A_{\mathrm{dev}} \cong 0.52$ degrees for the Simrad ES120-7 120-kHz transducer), about 57.9% of the unresolvable multiple targets were categorized as individual scatterers. On the other hand, due to the virtual elimination of scattering interference events, the two-frequency method (as defined in the preceding paragraph) greatly reduces the false acceptances to less than 30%, even when using broad limits on the discrepancy in off-axis angles as measured with the two frequencies [Fig. 1(b)].

In order to more closely match a typical EK500 survey configuration, the simulation was repeated with pulse lengths of 1.0 and 0.3 ms, quadrant separations of 79 and 42 mm, and angle sensitivities of 12.5 and 21, at 38 and 120 kHz, respectively, and runs of 1000 pings were evaluated at the default P_{dev} limits of four phase steps ($\alpha=\beta=0.92$ at 38 kHz and 0.52 degrees at 120 kHz. Under this common configuration, 75.4% of the multiple targets were falsely accepted at 38 kHz, 25.3% at 120 kHz and only 3.6% to 18.4% using the two-frequency method with angle-discrepancy criteria of 0.1 to 1.5 degrees, respectively [Fig. 1(c)].

In short, these simulations predict that the two-frequency method for rejecting unresolvable targets will provide a performance gain of at least fourfold over the single-frequency algorithm (Note that some of the performance gains are made through a reduction in pulse length at 120 kHz.) However, in order to effectively match a target in range and off-axis angles at two frequencies simultaneously, the positional information must be precise. Exactly how precise must be answered by the requirements of the investigation and is dependent upon the SNR (Table I).

To investigate this consideration, noise measurements were recorded under actual survey conditions while conducting a survey off the South Shetland Islands, Antarctica. The 38-, 120-, and 200-kHz transducers (Simrad ES38-12, ES120-7 and ES200-28) were hull-mounted in a steel blister at a depth of about 7 m, about 30 m aft-of-bow on the 105-m-long vessel (R/V YUZHMORGEOLOGIYA). Background noise power levels (P_n) were recorded by the three EK500 receivers, operating in passive mode, with sea conditions at Beaufort 5 while surveying at a ship speed of 10 kn.

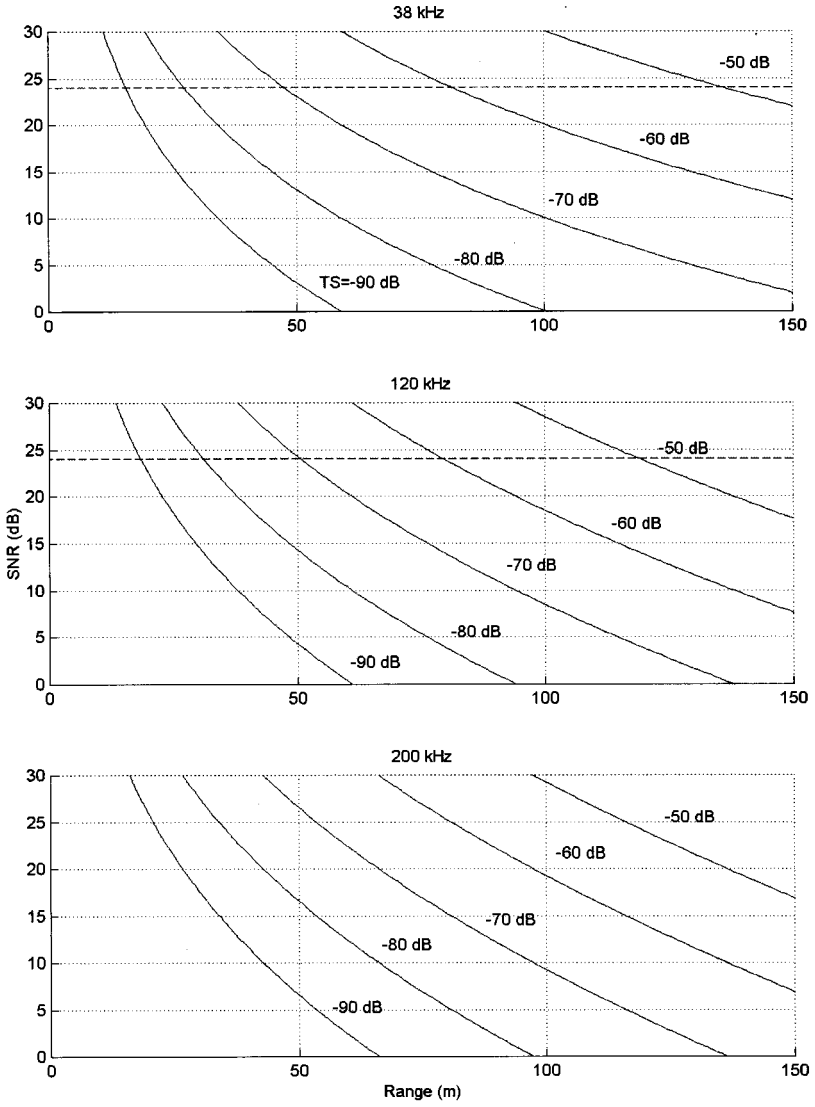


FIG. 2. Signal-to-noise ratio versus range for various TS values (solid curves). At the respective frequencies of 38, 120, and 200 kHz, the absorption coefficients (γ) were 0.01, 0.038, and 0.053 dB/m; the on-axis system gains (G) were 22.8, 24.7, and 28.0 dB; the wavelengths were 0.0382, 0.0121, and 0.0073 m; the noise powers (P_n) were measured to be -136.8, -147.0, and -148.6 dB *re*: 1 W; the pulse lengths (τ) were 1.0, 1.0, and 0.6 ms; and the transmit powers (P_t) were all 30 dB *re*: 1 W. For the EK500 echosounder operating with these conditions and parameters, maximum precision of the phase measurement is possible at a SNR of more than 24 dB (dashed line).

Under these survey conditions, the background power levels were $P_n = -136.8 \pm 0.3$ dB, -147.0 ± 0.3 dB, and -148.6 ± 0.1 dB *re*: 1 W, at 38, 120, and 200 kHz. Using these values of noise power and solving Eq. (6) for received power, the SNR was plotted versus range for various TS values (Fig. 2). Calculated from (15) and also plotted in Fig. 2 is the requisite SNR to potentially measure off-axis angles with the EK500 split-beam system to within one phase step (SNR$>$24 dB). Fortunately, Fig. 2 indicates that the maximum angular precision of the EK500 is not noise limited under typical survey conditions. However, the maximum range for this angular precision is strongly dependent upon the scattering strength of the target [Eq. (6)]. For example, considering only the constraints of noise, the off-axis angles of a euphausiid (TS~ -70 dB at 120 kHz) can be measured with maximal precision out to a range of about 50 m.

II. TANK EXPERIMENTS

Controlled experiments were conducted in a large test tank at the Institute of Maritime Technology in Simonstown, South Africa from 18 June to 26 July 1996. The tank (approximately 20 m long by 10 m wide by 10 m deep), contained fresh water at a temperature of 18 °C. An EK500 echosounder (firmware version 5.2) was configured with 38- and 120-kHz split-beam transducers (Table III). The transducers were mounted next to each other, 4 m deep and 6 m from one end of the tank, so as to project horizontally down the length of the tank.

A. TS measurement precision

1. On-axis measurements

In order to utilize the multi-frequency method for *in situ* TS measurements, it was first necessary to characterize the

TABLE III. Echosounder and transducer specifications (nominal).

Simrad EK500 (kHz)	38	120	200
Frequency (kHz)	37.878	119.047	200.000
Transducer model	ES38.12	ES120-7	ES200-28E
3 dB beamwidth (degrees)	12±1	7.1	7.0±1
Angle sensitivity (Λ)	12.5	21.0	[a]
Angular resolution (ζ) (degrees)	0.225	0.134	[a]
Range cell (m)	0.10	0.03	0.02
Pulse duration (τ) (ms)	1.0	0.3	0.6
Receiver bandwidth (kHz)	3.8	12.0	2.0

[a]Not applicable for single-beam transducer.

TABLE IV. Nominal on-axis TS measurements of standard spheres at 38 and 120 kHz in fresh water at a temperature of 18.9 °C. Also shown are (1) the differences in TS at these two frequencies; (2) the ratio of these differences to the sphere diameters; and (3) the slope of TS versus the logarithm of the product of the wave number and the spherical radius (a). All measurements were made with the same system gain settings. The TS measurements typically ranged at least ±0.5 dB (see Fig. 2).

Sphere diameter ($2a$)	23 mm Cu	30.05 mm Cu	38.1 mm WC	60 mm Cu
$TS_{120\ kHz}$ (dB)	−40.4	−36.3	−39.3	−33.3
$TS_{38\ kHz}$ (dB)	−48.0	−40.5	−42.3	−33.1
$TS_{120}-TS_{38\ kHz}$ (dB)	7.6	4.2	3.0	−0.2
$(TS_{120}-TS_{38\ kHz})$/Diameter (dB/m)	330.4	139.8	78.7	−3.3
$ka_{120\ kHz}$	5.8	7.6	9.6	15.1
$ka_{38\ kHz}$	1.8	2.4	3.0	4.8
Slope of TS vs Log(ka)	55.7	44.8	45.4	33.0

measurement precision of the split-beam system. First, TS measurements were made at both frequencies of four standard spheres [23.0, 30.05, and 60 mm copper (Cu), and 38.1 mm diameter tungsten carbide with 6% cobalt binder (WC)] (Table IV). In succession, each standard sphere was placed on the beam-axis approximately 10 m from the transducer. The spheres (Foote, 1983, 1990) were suspended by a monofilament line which was attached to the WC sphere by a monofilament knotted bag, and to each Cu sphere by a loop of monofilament nylon affixed into a single shallow bore.

At both 38 and 120 kHz, TS measurements of the four spheres exhibited variances typically greater than 0.5 dB (Table IV). For the three largest spheres, the differences in TS (120–38 kHz) decreased linearly versus increasing sphere diameter. For the 23-mm Cu sphere, which is approximately 60% smaller than the wavelength at 38 kHz (ap-

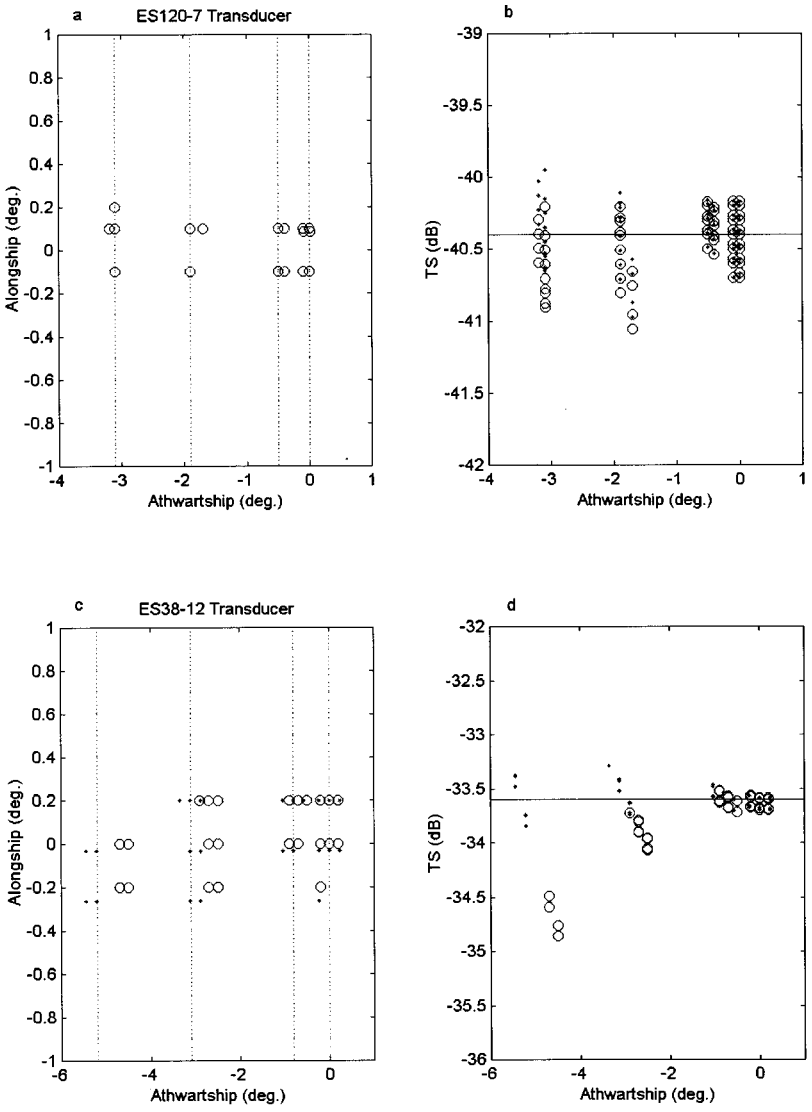


FIG. 3. Measurement of transducer angle sensitivity. For the ES120-7 transducer [(a) and (b)], a 23-mm Cu sphere was positioned at a range of 10.0 m and off-axis distances of 0, 87, 325, and 537 mm (0.0, −0.5, −1.9, and −3.1 degrees). From these TS measurements (○), the nominal value for this transducer ($\Lambda=21.0$) was shown to differ from value judged optimal ($\Lambda=20.5$, obtained by detrending the TS data with respect to off-axis angle) to within the angular resolution (≅0.134 degrees). The measurements were repeated (○) for the ES38-12 transducer [(c) and (d)] using a 60 mm Cu sphere at a range of 10.0 m and off-axis distances of 0, 147, 550, and 909 mm (0.0, −0.8, −3.1, and −5.2 degrees). The nominal value for this transducer ($\Lambda=12.5$) was shown to differ greatly from the value judged optimal ($\Lambda\cong10.8$). The lesser angular resolution at 38 kHz (≅0.225°) results in greater measurement variability. At both frequencies, the corrected angle sensitivity reduced the bias in off-axis TS measurements (•).

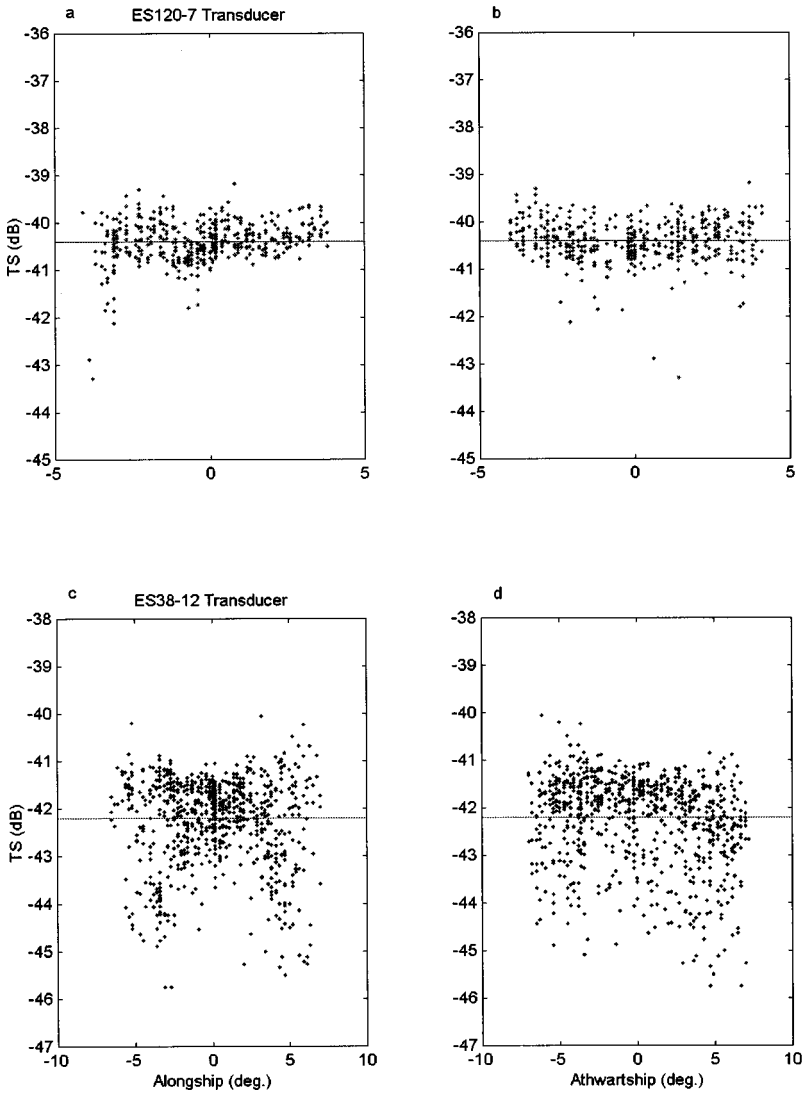


FIG. 4. Precision of TS measurements versus off-axis angle. In the test tank, the TS of a 23 mm Cu sphere was measured as it was moved, at a range of 10 m, through the beams of the ES120-7 transducer. Likewise, the TS of a 38.1 mm WC sphere was measured with the ES38-12 transducer. Plotted are the measured (•) and theoretical TS values (–).

prox. 39 mm), the TS difference versus the sphere diameter is more than three times greater than that for the others.

2. Off-axis measurements

The precision of off-axis measurements is primarily related to the accuracy of Λ, $B(\alpha,\beta)$ and ζ, and the SNR. The nominal angle sensitivity for the 120-kHz transducer was measured next by placing a 23-mm Cu sphere 10.0 m from the transducer and off-axis by distances of 0, 87, 325, and 537 mm (0.0, −0.5, −1.9, and −3.1 degrees). The measurements were repeated for the 38-kHz transducer using a 60-mm Cu sphere at a range of 10.0 m and off-axis distances of 0, 147, 550, and 909 mm (0.0, −0.8, −3.1, and −5.2 degrees) (Fig. 3).

For the ES120-7 transducer, the measured value for Λ (20.5) differed from the nominal value (21.0). For the ES38-12 transducer, the difference was much greater (measured $\Lambda \cong 10.8$ versus nominal $\Lambda = 12.5$). At both frequencies, the corrected angle sensitivity reduced the bias in off-axis TS measurements (Fig. 3). Note also that the decreased angular resolution at 38 kHz ($\cong$0.225 degrees) results in significantly greater measurement variability.

Using the measured angle sensitivities, the general precision of individual TS measurements was characterized by moving an individual sphere within the confines of the transducer 3-dB beamwidth. The TS values were recorded versus off-axis angles at 38 kHz using the 38.1-mm WC sphere and at 120 kHz using the 23-mm Cu sphere (Fig. 4).

As a 23-mm Cu sphere was moved through the beam of the ES120-7 transducer at a range of 10 m, the TS measured at 120 kHz ranged 4.1 dB (σ=0.4 dB) [Fig. 4(a) and (b)]. Similarly, while moving a 38.1-mm WC sphere through the beam of the ES38-12 transducer at 10-m range, the TS measurements ranged 5.7 dB (σ=1.0 dB). At 38 kHz, the decreased measurement precision is due to the lower range and angular resolutions, coupled with a possible mismatch between modeled and actual beam-patterns [Fig. 4(c) and (d)].

B. Multiple target rejection

1. Single-frequency algorithm

The effectiveness of the EK500 algorithm for rejecting multiple targets was evaluated for a benchmark by moving two spheres randomly within the same resolution volume (Fig. 5). At 38 kHz, the TS detections were recorded while two 60-mm Cu spheres (each nominal TS = −33.6 dB) were

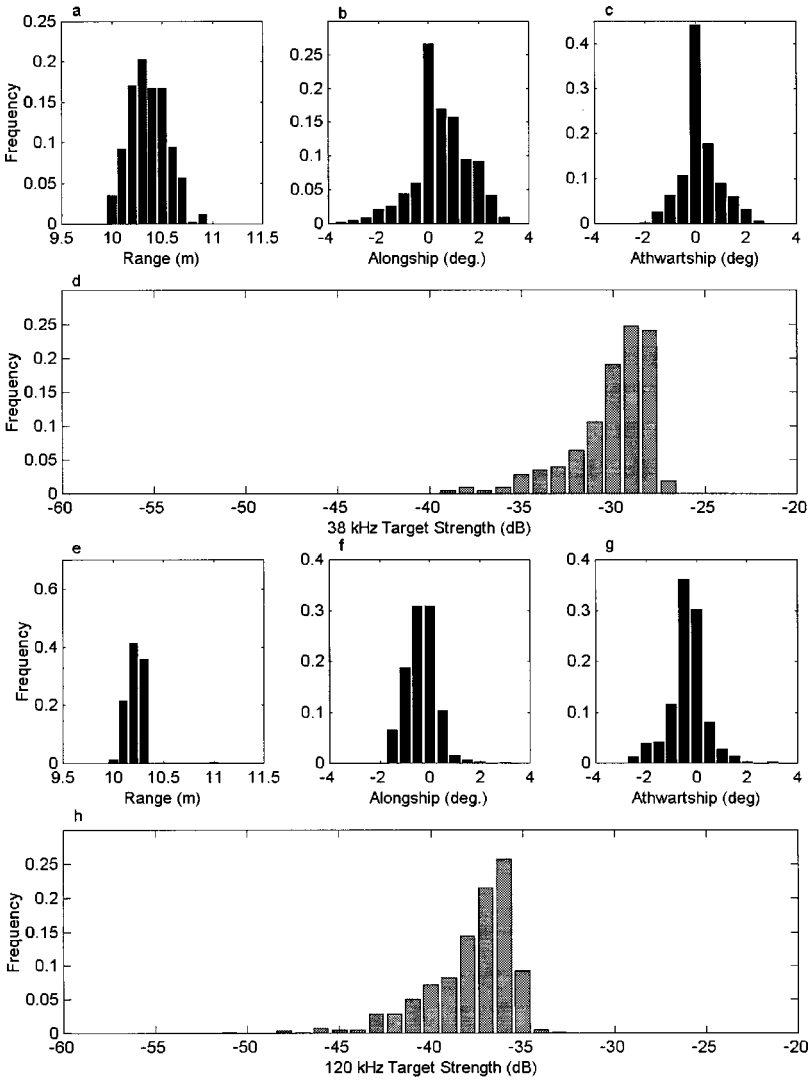


FIG. 5. Single target detections at 38 and 120 kHz from a field of two spheres. At 38 kHz [(a)–(d)], two 60-mm Cu spheres (each −33.6 dB; Foote, 1983) were moved randomly ±2.17 degrees off-axis (±0.380 m) and within a single range resolution volume (10 to 10.75 m). At 120 kHz [(e)–(h)], two 23-mm Cu spheres (each −40.4 dB) were moved randomly ±2.17 degrees off-axis (±0.380 m) and within a single range resolution volume (10 to 10.225 m).

randomly moved ±2.17 degrees off-axis (±0.380 m) and in range from 10 to 10.75 m. At 120 kHz, two 23-mm Cu spheres (each nominally −40.4 dB) were randomly moved ±2.17 degrees off-axis (±0.380 m) and in range from 10 to 10.225 m.

At 38 kHz, the EK500 falsely misinterpreted two acoustically unresolvable 60.0-mm Cu spheres (TS=−33.6 dB; Foote, 1983) as individuals in 35% of the 1200 pings (Fig. 5). Resulting from interference effects, the TS distribution is shifted upward (mean=−29.6 dB), ranges 12.2 dB (σ=2.2 dB), and has minimum values as low as −39.1 dB. Similarly at 120 kHz, the two 23.0 mm Cu targets (TS=−40.4 dB) were falsely misinterpreted as individuals in 40% of the 1500 pings. The distribution of measurements is negatively skewed (mean=−37.3 dB), ranges 27 dB (σ=2.5 dB), and has minimum values as low as −60 dB.

2. Multiple-frequency method

The effectiveness of matching TS detections at multiple frequencies for rejecting unresolvable echoes was then investigated. First it was necessary to derive tolerances for matching target positions at two or more frequencies. To do so, measurements of a single 23-mm Cu sphere (range, off-axis angle, and TS) were made simultaneously at 38 and 120 kHz. At a nominal range of 10 m, the sphere was moved horizontally and randomly inside the confines of a box 1.1 by 0.76 m (Fig. 6).

With the two transducers mounted adjacently (Fig. 7),

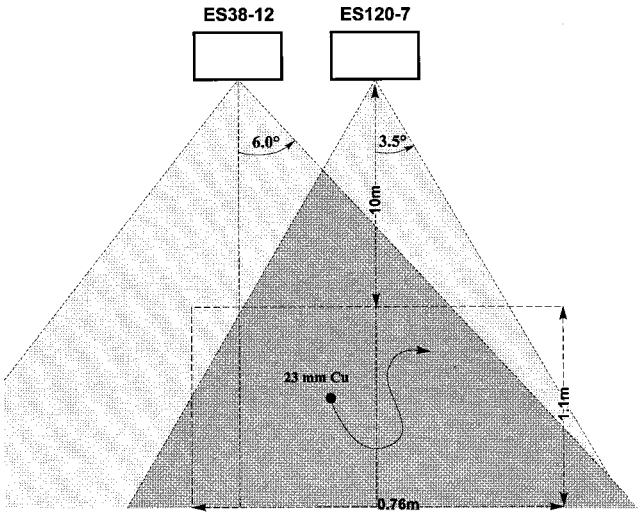


FIG. 6. Simultaneous TS measurements of a 23-mm Cu sphere at 38 and 120 kHz. At a nominal range of 10 m, the sphere was moved randomly inside the confines of a box 1.1 by 0.76 m^2. The drawing is not to scale.

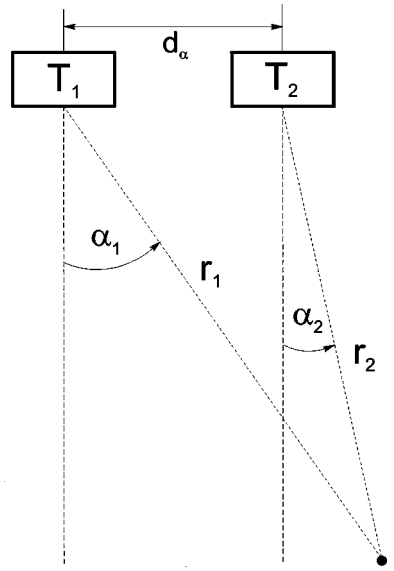


FIG. 7. Matching target detections from two adjacent transducers. The off-axis angles (α_1 and α_2) and the slant ranges (r_1 and r_2) are interrelated by the separation of the parallel transducer axes (d_α).

beam axes parallel, α, β, and r are related by the separation of the beam axes in the alongship and athwartship planes (d_α, d_β):

$$\alpha_2 = \sin^{-1}\left(\frac{r_1 \sin\alpha_1 - d_\alpha}{r_2}\right)$$

and (17)

$$\beta_2 = \sin^{-1}\left(\frac{r_1 \sin\beta_1 - d_\beta}{r_2}\right).$$

The angular discrepancy between the two systems [theoretical α_{120} or β_{120} (calculated from $d_{38\text{-}120}$ and α_{38} or β_{38}) minus measured α_{120} or β_{120}, respectively] was ± 1.5 degrees (>95% confidence). This can be attributed to a combination of (1) the angular resolutions of the two systems; (2) some unknown rotational angles inherent in the transducer mounting or beam-pattern; (3) noise; and (4) the polychromatic transfer functions or scattering complexity of the individual scatterers.

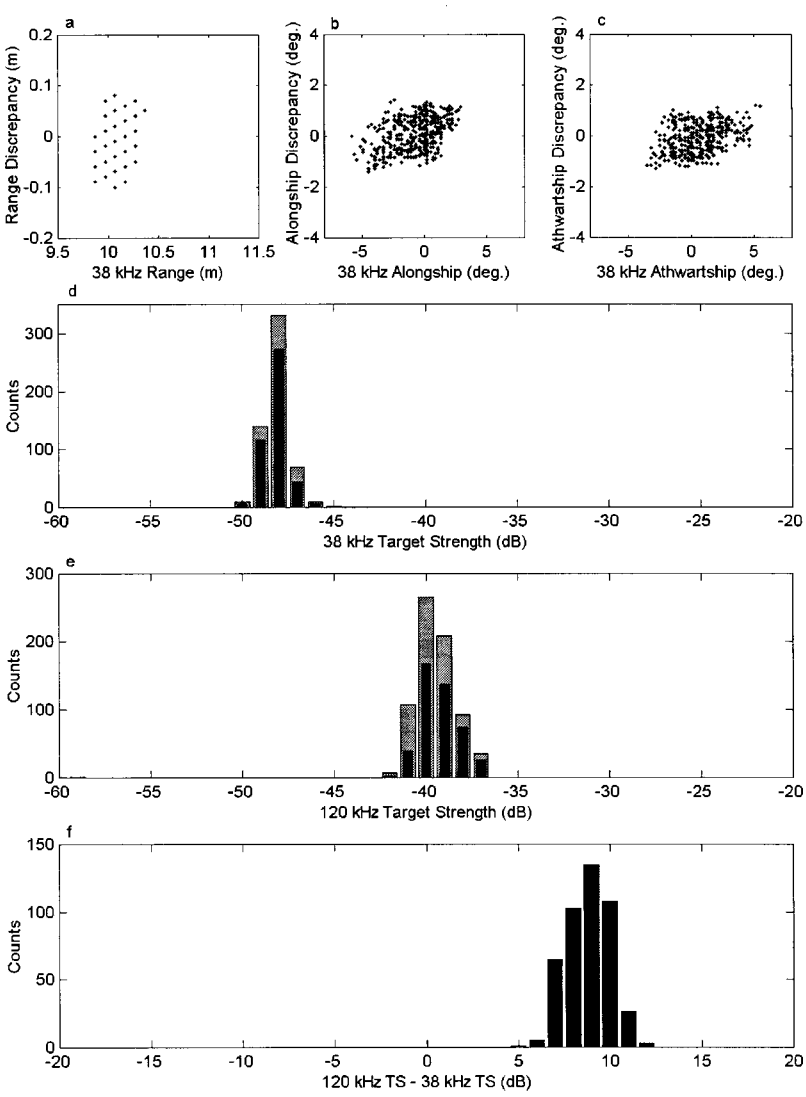


FIG. 8. The TS measurements of a single 23.0-mm Cu sphere at a range of 10 m, moved randomly through the beams of juxtaposed 38 and 120 kHz transducers. From 1000 pings, the single-frequency algorithm accepted 560 and 717 measurements at 38 and 120 kHz, respectively (gray bars). Of these, the two-frequency method accepted 440 measurements; a reduction of 21% to 39% (black bars).

Due to the differing bandwidths and sampling rates of the two systems and a small offset of the two transducer faces, target ranges at 38 kHz (r_{38}) were consistently about one sample range ($\cong$0.1 m) greater than the more accurate range estimates (r_{120}) of the 120-kHz system (sample range $\cong$0.03 m) and both ranges were slightly greater than the actual ranges. In order to compensate for the receiver delay, measurements of target ranges were reduced by three sample ranges or by 0.3 m at 38 kHz and 0.09 m at 120 kHz and then a residual mean offset of 0.07 m was subtracted from the 38-kHz ranges (Figs. 8, 10, and 12). After applying these corrections, the variation in range discrepancy [theoretical r_{120} (calculated from $d_{38\text{-}120}$ and measured r_{38}) minus measured r_{120}] for a single 23-mm Cu sphere was $\pm$0.1 m (>95% confidence). (Note, the EK500 firmware V5.2 and V5.3 subtracts three sample ranges and interpolates between range cells for the purpose of transmission loss compensation only—not range).

The tolerances for spatial matching were then used in the multi-frequency method for rejecting multiple targets. Specifically, targets were only accepted as individual scatterers if the single-target detections at each frequency were matched to within $\pm$1 range cell at 38 kHz (0.1 m) and to an off-axis angular discrepancy of 1.5 degrees. The target ranges could be matched this closely because the filter averaging times ($\cong$1/bandwidth), in terms of range cells, are very small and virtually identical at both frequencies (1/3.8 kHz=0.263 ms/133 μs/sample=1.9725 samples at 38 kHz and 1/12.0 kHz=0.08 ms/40 μs/sample=2 samples at 120 kHz). For a single 23-mm Cu sphere moved randomly inside the two partially overlapping beams, single-target detections with the single-frequency algorithm totaled 560 and 717 out of 1000 pings at 38 and 120 kHz, respectively (Fig. 8). In contrast, the two-frequency method accepted only 440 measurements or 21% to 39% fewer than the single-frequency method with no bias or change in variance (Fig. 8).

For both the single- and multiple-frequency algorithms, the ranges of TS measurements were greater than 5 dB. It is also important to note that the variance in TS differences [var$(TS_{120}-TS_{38})$] is worse than either of the single-frequency measurements due to the compounding of measurement errors; it is expected to be the sum of the variances of the single-frequency measurements plus twice the covariance (Rice, 1988).

The TS measurements were also made of three spheres residing within 1.47 and 4.89 range cells at 38 and 120 kHz, respectively (Fig. 9). The 23-mm Cu sphere was fixed into position on the beam axis of the 120-kHz transducer at a range of 10 m while a 30.05-mm Cu and a 38.1-mm WC sphere were moved randomly inside the confines of the 1.1 by 0.76 m box. Two-frequency TS measurements were again recorded.

Theoretically, a single target could be accurately detected in this configuration, but the probability of a single sphere residing within one 38-kHz resolution volume is low (0.175 to 0.225 with 95% confidence from Monte Carlo simulation of 1000 runs of 1000 pings) and delineation of more than one sphere per ping is physically impossible. Nonetheless, out of 1000 pings, detections with the single-frequency algorithm totaled 329 at 38 kHz and 1059 at 120 kHz (some multiple detections per ping).

Applying the two-frequency method, 89% and 96% of the single-frequency detections were rejected as multiple-scatterers at 38 and 120 kHz, respectively (Fig. 10). Of the remaining 38 TS measurements, there was no more than one target detection in any single ping. The histogram of TS differences exhibits a main peak corresponding to the 30.05-mm Cu sphere and the 38.1 mm WC sphere and a smaller peak corresponding to the 23-mm Cu sphere. Additionally there are two small peaks at plus and minus 1 dB, probably due to a residual of multiple targets accepted [0.6% (38 kHz) to 1.8% (120 kHz) of single-frequency detections] with the Version 5.2 default detection criteria (Table II).

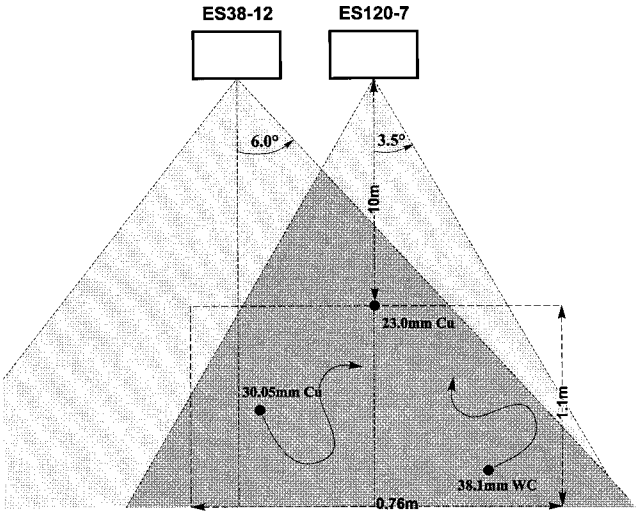


FIG. 9. Simultaneous TS measurements of a 23-mm Cu sphere were made at 38 and 120 kHz while a 30.05-mm Cu and a 38.1-mm WC sphere were moved randomly inside the confines of a box 1.1 by 0.76 m^2. The drawing is not to scale.

III. FIELD EXPERIMENTS

In situ target strength measurements were collected off the western side of the Antarctic Peninsula from mid-January through mid-February 1997. The shipboard echosounder system (Simrad EK500; firmware version 5.2) was configured with 38- and 120-2 kHz split-beam transducers (different than those used in the tank experiments, but the same models) and a 200-kHz single-beam transducer (Table III). The down-looking transducers were mounted on the hull of the ship, in a row (from fore-to-aft: ES120-7, ES38-12, ES200-28), approximately 7 m deep. The system was calibrated using a 38.1-mm WC sphere (Foote, 1990), before and after the survey at Ezcura Inlet, King George Island. Acoustic transects were conducted for 14 consecutive days in an area around Elephant Island (AMLR, 1997). Approximately every 30 km, acoustic targets were sampled with a 2.5-m^2 Isaacs-Kidd midwater trawl (IKMT) fitted with a 505-μm mesh net (Devereaux, 1953).

A total of 105 IKMT tows were conducted concurrently with collections of acoustic target strength. From the samples, 73 zooplankton and nekton species and categories

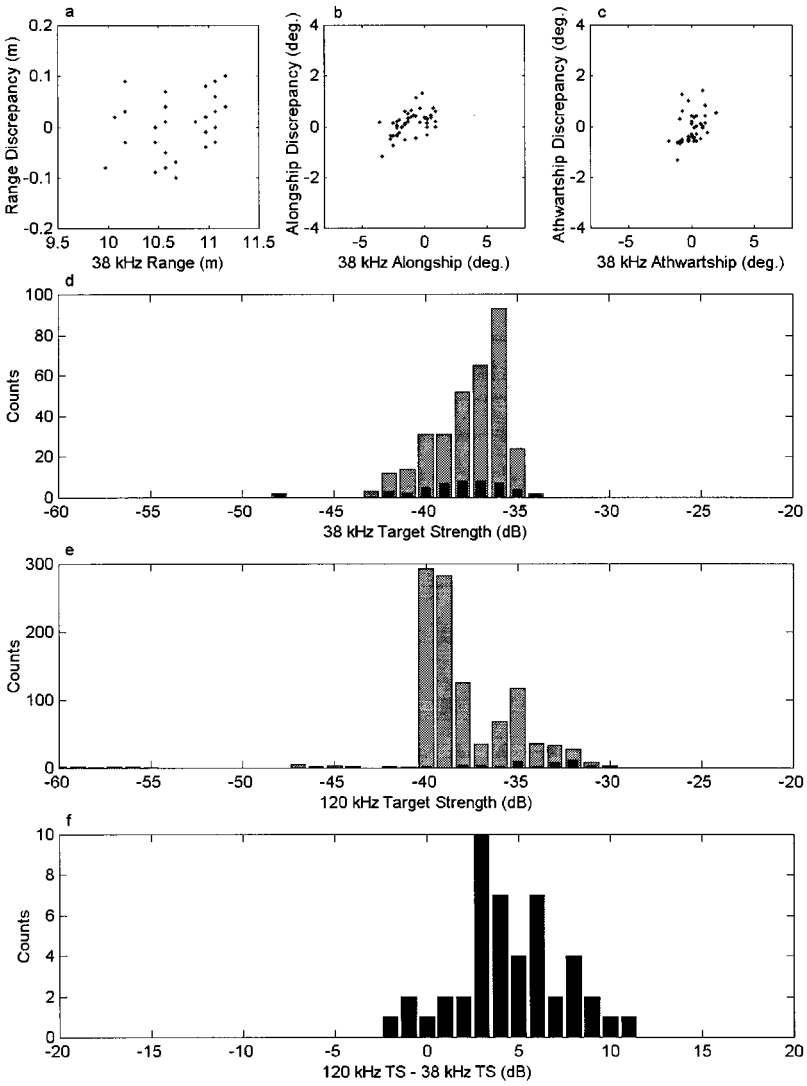


FIG. 10. Testing the two-frequency method for multiple target rejection. A 23-mm Cu sphere was positioned on the axis of the 120-kHz transducer at a range of 10 m while two other spheres (30.05 mm Cu and 38.1 mm WC) were randomly moved around the sphere within 1.47 and 4.89 range cells at 38 and 120 kHz, respectively. The probability of accurately detecting an individual scatterer in this proximity is very low. From 1000 pings, detections with the single-frequency algorithm totaled 329 at 38 kHz and 1059 at 120 kHz (gray bars). In sharp contrast, the two-frequency method passed only 38 measurements, rejecting 89% to 96% as multiple-scatterers (black bars).

were identified, including crustacean invertebrates, gelatinous tunicates, chaetognaths, and myctophid fishes. The four most numerically abundant taxa, caught in over 90% of the tows, were (1) copepods; (2) a pelagic tunicate (*Salpa thompsoni*, with length modes of individual animals at 10, 25, and 45 mm); (3) two euphausiids [*Thysanoessa macrura*, less than 20 mm mean length and *Euphausia superba* (Antarctic krill) with length modes at 26, 37, and 50 mm]; and (4) an amphipod (*Themisto gaudichaudii*, less than 10 mm mean length). The taxa with the largest animals included euphausiids (*T. macrura* and *E. superba*), and two species of myctophids (*Electrona carlsbergi*, range 70–90 mm; and *Electrona antarctica*, range 40–110 mm). The myctophids were caught in approximately 10% of the tows, although their distribution appeared to be much more widespread from an examination of the echograms. Their under-representation in the IKMT catches may be explained by the ability of these fish to avoid the relatively small net.

Before the multi-frequency method could be applied to the field measurements of *in situ* TS, the minimum detection ranges were determined. For simultaneous detections at two (38 and 120 kHz) and three frequencies (38, 120, and 200 kHz), the minimum ranges ($r\min_{38\text{-}120}$ and $r\min_{200\text{-}120}$) are functions of the transducer separations ($d_{38\text{-}120}$ and $d_{200\text{-}120}$) and the maximum allowable off-axis angles for each transducer ($\phi_{38}, \phi_{120}, \phi_{200}$):

$$r\min_{38\text{-}120} = \left(\frac{d_{38\text{-}120}/\tan(\phi_{38})}{1+(\tan(\phi_{120})/\tan(\phi_{38}))} \right)$$

and (18)

$$r\min_{200\text{-}120} = \left(\frac{d_{200\text{-}120}/\tan(\phi_{200})}{1+(\tan(\phi_{120})/\tan(\phi_{200}))} \right).$$

For the shipboard transducer mounting configuration (Fig. 11; $d_{120\text{-}38}=0.4425$ m and $d_{38\text{-}200}=0.4425$ m, and targets accepted within the 3-dB beamwidths ($2\phi_{38}\cong 12$ degrees, $2\phi_{120}\cong 2\phi_{200}\cong 7$ degrees), $r\min_{38\text{-}120}=2.7$ m and $r\min_{200\text{-}120}=7.2$ m.

When moving from a test tank to the field, the measurements of range and off-axis angle may be affected by increases in the number of noise sources and target complexity. However, if the condition of far-field operation is imposed (with respect to the apertures of both the transducer and the targets), the receiver bandpass filters (Table III) could be expected to reduce the measurement uncertainty due to noise and the complex nature of real scatterers. Therefore, consistent with the tank experiments, the single-target

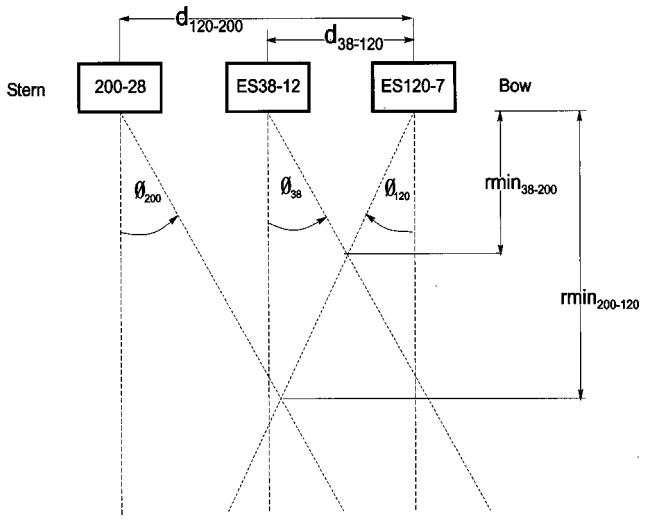


FIG. 11. Minimum ranges for TS matches using multiple transducers. The minimum ranges for simultaneous detections at two- and three-frequency systems ($r\text{min}_{38\text{-}120}$ and $r\text{min}_{200\text{-}120}$) are functions of the maximum allowable off-axis angles (ϕ) and the horizontal separations between beam axes ($d_{38\text{-}120}$ and $d_{200\text{-}120}$, respectively).

detections at 38 and 120 kHz were matched to within ±1 range cell at 38 kHz (0.1 m) and constrained to an off-axis angular discrepancy of 1.5 degrees. Once again, 82% to 96% of the single-frequency single-target detections were rejected as multiple targets (Fig. 12).

The *in situ* TS measurements were plotted versus range for one-frequency [Fig. 13(a) and (b)] and two-frequency detections [Fig. 13(c) and (d)]. With the two-frequency method, a large proportion of the unresolvable targets were rejected at large ranges. This was expected because the insonified volumes increase with increasing range, and the probabilities of resolving individual animals decreases accordingly. Of note, however, is the preferential rejection of smaller targets as described by the SNR required (24 dB) to measure off-axis angles to within one quantization step. Approximately 87% to 98% of the two-frequency matches were above this SNR at 38 and 120 kHz, respectively.

Corresponding TS values at 200 kHz were extracted from single-beam data by matching range bins to within ±1 range cell at 38 kHz. Consistent with the firmware implementation at the other two frequencies (Simrad, 1996), the

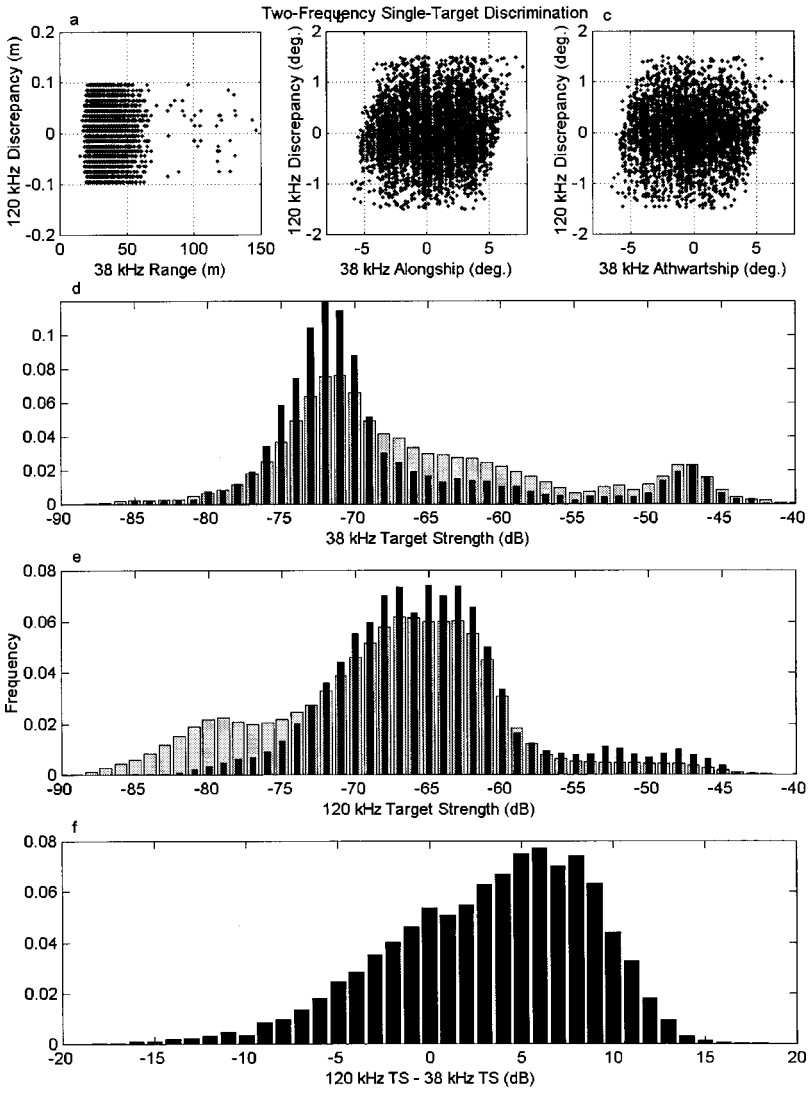


FIG. 12. *In situ* TS measurements of individual Antarctic zooplankton and nekton. Total detections using a single-frequency method (gray bars) were 40 391 at 38 kHz (d) and 173 512 at 120 kHz. (e) Using the two-frequency method, these numbers were reduced by 82% to 96% as only 7248 individual scatterers were detected (black bars).

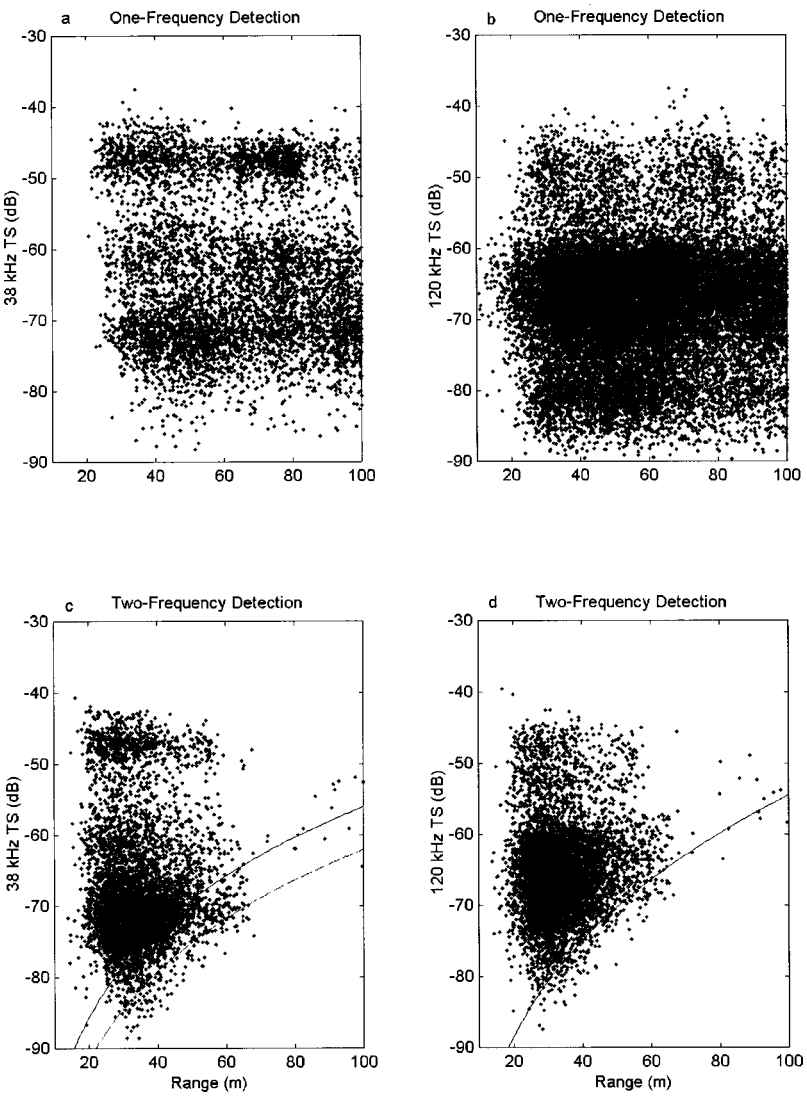


FIG. 13. Scatter plots of *in situ* TS versus range for one [(a) and (b)] and two-frequency detections [(c) and (d)]. With the two-frequency method, most unresolvable targets are rejected at large ranges where the insonified volumes are increasingly larger and the probability of resolving an individual animal is increasingly lower [(c) and (d)]. Also apparent is the rejection of smaller targets with a SNR smaller than that required (24 dB) to measure off-axis angles to within one quantization step (solid curve). At 38 and 120 kHz, 87.3% and 98.0% of the two-frequency matches were above this 95% confidence interval, respectively. At 38 kHz, 98.6% of the matches were above the SNR required (18 dB) to measure off-axis angles to within two phase quantization steps (dashed curve).

approximated one-way beam pattern $[\hat{B}(\alpha,\beta)]$ of the circularly symmetric transducer was

$$\hat{B}(\alpha,\beta) = -3\left[\left(\frac{\alpha}{\phi/2}\right)^2 + \left(\frac{\beta}{\phi/2}\right)^2 - 0.18\left(\frac{\alpha}{\phi/2}\right)^2\left(\frac{\beta}{\phi/2}\right)^2\right], \tag{19}$$

where ϕ is the 3-dB beam width of the transducer. Beampattern effects were thus removed from the resulting 200 kHz single-beam data by using the split-beam positional information of the 120-kHz transducer (Fig. 14).

As is frequently the case in fisheries acoustics, the net samples were suspected of bias (especially for nekton) and indisputable matches to the *in situ* TS measurements were impossible. Therefore, the backscattering taxa must be inferred from a combination of catch data, modes in the TS distributions, and theoretical expectations of scatterer reflectivity as a function of acoustic wavelength and animal size (Chu *et al.*, 1992).

Euphausiids were expected to exhibit Rayleigh scattering at 38 kHz and geometric scattering at 120 and 200 kHz (Demer and Martin, 1995). As such, mean TS values were expected to increase with frequency and the differences between TS measurements of euphausiids at 38 kHz and those at 120 and 200 kHz were expected to be larger and have greater variance than for animals scattering in the geometric regime at all three frequencies (e.g., fish). In the case of fish, TS was expected to decrease or increase versus frequency, depending on the existence or lack of swim-bladders, respectively (MacLennan and Simmonds, 1992).

Prior to the application of the multi-frequency method, three modes were apparent in the 120-kHz TS distribution (Figs. 12 and 13): a primary mode at about −65 dB, a secondary mode at approximately −79 dB, and a tertiary mode at about −50 dB. When selecting only those targets detected by 38 and 120 kHz, the secondary mode, possibly due to individual *S. thompsoni* or small euphausiids, was eliminated by thresholding (Foote, 1991; MacLennan and Simmonds, 1992) at 38 kHz.

After selecting only those targets which were detected simultaneously at three frequencies (Fig. 14), three modes are apparent in the 38-kHz TS distributions (Table V). Mean TS values for mode 1 were 5 to 6 dB lower at 38 kHz relative to the higher frequencies, indicating small scatterers relative to the largest wavelength (39 mm). For mode 2, the mean TS values were virtually identical, suggesting a large scatterer relative to all three wavelengths. Mode 3 averaged 3

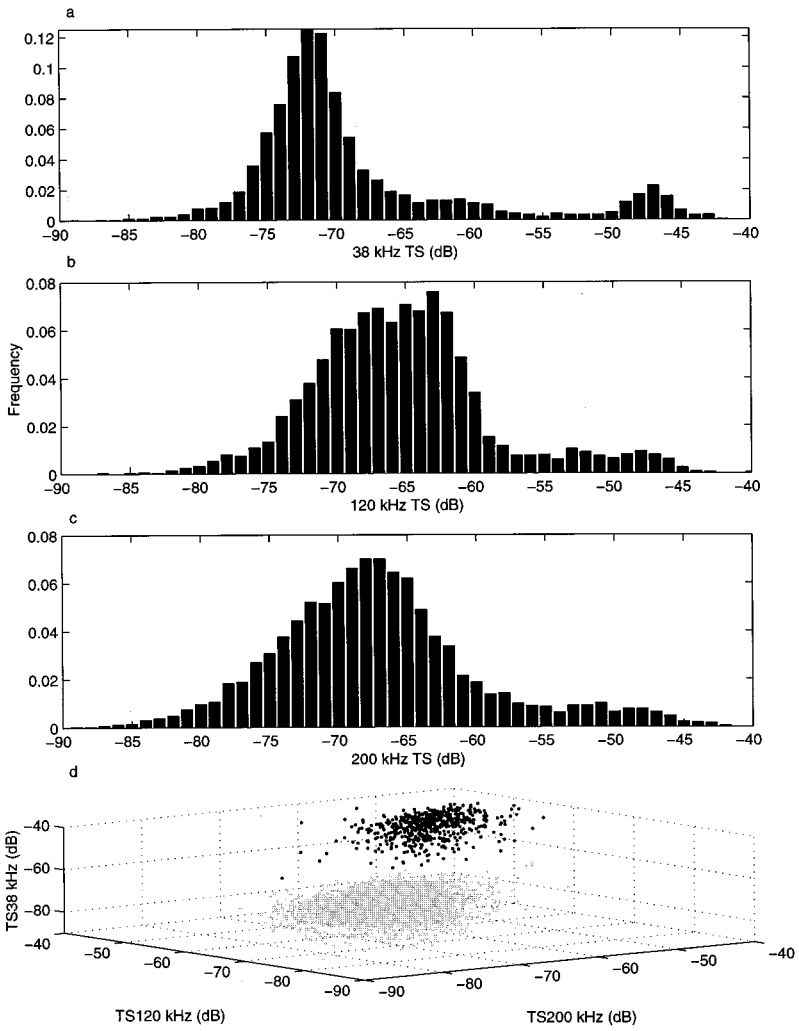


FIG. 14. *In situ* TS of Antarctic scatterers at 38, 120, and 200 kHz. Matches of single-target detections at all three frequencies totaled 5690. The 38 kHz histogram [(a)] contains three modes: $-85 \leqslant$ mode $1 < -65$ dB; $-65 \leqslant$ mode $2 < -55$ dB; and $-55 \leqslant$ mode $3 < -40$ dB. Mode 1 (black dots) and mode 3 (gray dots) are clearly distinguishable by the three-frequency TS characterization [(d)].

dB higher TS values at 38 kHz relative to 120 and 200 kHz, possibly due to scattering from a swim-bladder. The predominant scattering species, which may be attributed to these modes, are Antarctic krill (*E. superba*) for mode 1 and *Electrona spp*. for modes 2 and 3. Improved correlations may be obtained by matching TS distributions and catch data in smaller increments of time and space.

IV. DISCUSSION

The application of echo integration methods for accurately measuring areal density of pelagic scatterers should involve simultaneous and accurate measurements of volume backscattering strength (S_v) and TS of the individual scatterers. When measuring TS *in situ*, it is necessary to eliminate measurements when multiple animals reside in the same sample volume. This constraint is particularly important when making measurements of macro zooplankton (e.g., euphausiids) or small pelagic fish (e.g., myctophids) which frequently aggregate in high density swarms and schools.

In situ TS measurements of animals at the periphery of conspecific aggregations may be more representative of animals encountered during the survey than TS distributions generated from theoretical models or controlled experiments. However, it is possible that scattering from solitary animals is not representative of that from more densely aggregated animals. Furthermore, thresholding can bias the measure-

TABLE V. *In situ* TS measurements at three frequencies. Tabulated are the means and standard deviations (σ) for three modes in the 38-kHz histogram and the same for target matches at 120 and 200 kHz. Mean TS values for mode 1 were 5 to 6 dB lower at 38 kHz relative to the higher frequencies, indicating scatterers with small backscattering areas relative to the largest wavelength (39 mm). For mode 2, the mean TS values were virtually identical, suggesting a large scatterer relative to the wavelengths. Mode 3 averaged 3 dB higher at 38 kHz relative to 120 and 200 kHz, possibly due to scattering from a swimbladder.

Frequency (kHz)	$-85 \leqslant$ mode $1 < -65$ dB	$-65 \leqslant$ mode $2 < -55$ dB	$-55 \leqslant$ mode $3 < -40$ dB
38	−71.1 (3.1)	−60.5 (2.5)	−47.2 (2.7)
120	−65.0 (4.7)	−61.4 (4.7)	−50.0 (5.0)
200	−66.1 (5.3)	−60.2 (5.6)	−50.1 (5.4)

ments if inappropriate frequencies are chosen for the scattering taxa of interest; this is also the case for measurements of S_v.

This multi-frequency extension to the single-target detection algorithm proved very efficient at excluding unresolvable targets, relative to the single-frequency algorithm, in simulations, test tank experiments, and field studies. In one Antarctic survey, results suggest that the multi-frequency method provided a 98.2% to 99.4% improvement over the single-frequency method. There was some indication of measurement bias due to thresholding, but this can be minimized via optimal selection of frequencies for the scatterers of interest. By considering the known or expected distributions of animal size, shape, orientation, and acoustic impedance, adequate signal-to-noise ratios must be ensured for the measurements at all frequencies employed.

Simultaneous TS measurements at three frequencies (using two split-beam and one single-beam transducers) were made possible by utilizing positional information from a split-beam system to beam-compensate the data from an adjacent single-beam system. Differences in the resulting multiple-frequency TS measurements provided information about the identity of constituents in a mixed-species assembly. It should be noted, however, that the imprecision in single-frequency TS measurements is compounded in the estimation of TS and S_v differences. This measurement uncertainty, thresholding effects, and especially the sparse spectral sampling will limit the power of this taxa identification method.

The ultimate utilities of both the multi-frequency method for *in situ* TS measurements and the TS differences technique for taxa identification depend upon the uncertainty in the basic measurements. For example, the filtering efficiency of target matches at multiple-frequencies depends upon: (1) precise estimates of $\hat{\alpha}$, $\hat{\beta}$, and $\hat{r}$; (2) parallel beam axes or an accurate transform function; (3) low noise; and (4) low target complexity and narrow receiver bandwidth. This is illustrated through the simulation results (Fig. 1), which shows that the multi-frequency method is most efficient if the angular discrepancy between the two frequency detections is tightly constrained.

The angular resolution is a function of the phase quantization of the echosounder and the angle sensitivity of the transducer [Eq. (3)]. In turn, the angle sensitivity is dependent upon the transducer shading (Wilson, 1988; Foote, 1990) and the acoustic wavelength. Therefore, tolerances in the transducer manufacturing may produce inequities between the theoretical d_{eff} and the realized effective spacing. In addition, a change in the sound speed will result in the same percentage change in λ and Λ. Judging from these studies, the precision of the angular measurements (α and β) due to variances in d_{eff} and wavelength were approximately ± 0.26 degrees for the Simrad ES38-12 and ± 0.14 degrees for the ES120-7.

Optimally, the transducers should be mounted so the beams are projecting parallel to one another. If the actual angle discrepancies change as a function of range, the transducer geometry in Fig. 7 is not accurate and Eq. (17) must be appropriately modified (see Fig. 15). One method is to determine a spatial transform equation for each of the transducer pairs by moving a target within the overlapping beams and recording the associated spatial information (α, β, and r).

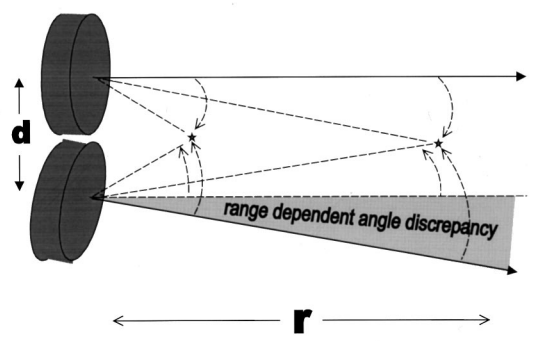


FIG. 15. Angle discrepancy as a function of range for nonparallel beams. For the multiple-frequency method, the transducers should optimally be mounted so the beams are projecting parallel to one another. If the beams are not parallel in both the alongship and athwartship planes, it is necessary to derive a spatial transform equation for each of the transducer pairs using the target positional data (α, β, and r) collected while moving an object within the overlapping beams.

Using the EK500, the maximum angular precision is noise limited at SNR less than 24 dB [Eqs. (6) and (17)]. As seen in Fig. 13, nearly all of the matching field measurements had larger SNRs and were therefore not noise limited. It is also important to note that the maximum angle discrepancy in the two-frequency method was 1.5 degrees ($\sim$6.7 phase steps for the ES38-12), or approximately seven times the angle discrepancy which could be attributed to noise. This implies that precise angular matching, and consequently the efficacy of the multi-frequency TS detection method, is limited by some process other than noise.

The multi-frequency method has the potential to greatly improve the accuracy and precision of *in situ* TS measurements by efficiently filtering multiple-target interference events which tend to pass the single-frequency algorithm. The realized effectiveness of the method is primarily dependent upon the minimum angular discrepancy criteria that can be imposed without filtering valid single targets. In these studies, the minimum angular discrepancy criteria was not limited by echosounder phase quantization, parallelism of the transducer beams, nor ambient noise. Therefore, by deduction, the method is likely limited by target complexity and the finite receiver bandwidths.

Echosounder designs and deployment configurations include many competing considerations, one of which is the effect of narrow-bandwidth receiver filters on both the SNR and the measured echo amplitude and phase. In general, near-optimal SNR can be achieved with receiver bandpass filters of width$=2/\tau$. The tradeoff is an echo signal with amplitude and phase that are functions of the bandpass filter—especially for narrower bandwidth filters. For pulses of 1 and 0.3 ms, the optimal filter bandwidths are 2 and 6.7 kHz, respectively. Unfortunately, the echosounder receiver bandwidths are user selectable to be either 1% or 10% of the transmit frequency. Because TS measurements were the priority of these experiments, and because the EK500 uses measurements of peak echo amplitude to estimate TS, ''wide''

bandpass filters ($BW = 0.1f$) were chosen. Consequently, the receiver bandwidths were 79% to 90% wider than optimal from a SNR perspective.

With appreciable bandwidth, narrow-band assumptions are not necessarily valid and the received signal $[p(t)]$ is derived from the complex transmit signal $[S(\omega)]$, the polychromatic backscattering amplitude of the acoustic scatterer $[F(\omega)]$, and the frequency response function of the receiver $[H(\omega)]$:

$$p(t) = \mathrm{FT}^{-1}(\mathrm{SFH}), \tag{20}$$

where the Fourier transform operator FT^{-1} transforms the argument from the frequency to the time domains (Foote, 1983). Because scattering is a stochastic process, the amplitude and phase of $p(t)$ can vary over time and frequency. This variability appears to be the limiting factor in the efficiency of the multi-frequency TS detection method. It also has ramifications for increased uncertainty in beam compensation of TS measurements in a split-beam system. Moreover, with the increasing development and usage of broadband acoustic systems (e.g., Simmonds *et al.*, 1996) for the purposes of TS measurements (e.g., Martin *et al.*, 1996), scatterer classification (e.g., Zakharia *et al.*, 1996) and also echo-integration (e.g., Thompson and Love, 1996), additional constraints due to the stochastic nature of scattering may be realized.

A very promising extension of this multi-frequency TS measurement method involves the synchronized usage of three or more collocated echosounders, at least two split-beam systems and another with very high range resolution. For example, if simultaneous transmissions were made at 38, 120, and 200 kHz with pulse lengths of 1.0, 0.3, and 0.06 ms, respectively, echoes are only accepted as individual targets if (1) the scatter is coherent at both 38 and 120 kHz and (2) the 200-kHz system ($c\tau/2 \cong 0.045$ m) does not resolve more than one target within the coincident sampling volumes. In this way, improvements are made to both the phase and range discriminators for multiple-target rejection.

ACKNOWLEDGMENTS

This research was supported by the U.S. Antarctic Marine Living Resources Program (AMLR) and the SFRI. Special thanks go to Manuel Barange, SFRI, and Merrick Whittle, IMT, for graciously funding and accommodating the test tank experiments, and to Valerie Loeb, Moss Landing Marine Laboratories, and Wesley Armstrong, AMLR, for summarizing the net sampling data. Thanks go to Tim Stanton, Woods Hole Oceanographic Institution, for advice and encouragement; to George Watters, International Tropical Tuna Commission, for discussions on noise statistics; and to three anonymous reviewers for constructive critiques.

AMLR (**1997**). *1996/97 Field Season Report: Objectives, Accomplishments, and Tentative Conclusions*, edited by J. Rosenberg, SWFSC Admin. Rep. LJ-97-09.

Bodholt, H., Nes, H., and Solli, H. (**1983**). *A new echo-sounder system for fish abundance estimation and fishery research*, ICES Council Meeting 1988 (collected papers) (ICES, Copenhagen, Denmark).

Bodholt, H. (**1990**). *Fish density derived from echo-integration and in-situ target strength measurements*, ICES Council Meeting 1990 (collected papers) (ICES, Copenhagen, Denmark).

Bodholt, H. (**1991**). ''Split-beam transducer for target strength measurement,'' in *Scandinavian Cooperation Meeting in Acoustics XIII*, edited by H. Hobæk, Department of Physics Sci./Tech. Rep. 227 (University of Bergen, Bergen, Norway), p. 73.

Bodholt, H., and Solli, H. (**1992**). ''Application of the split-beam technique for *in-situ* target strength measurements,'' World Fisheries Congress, Athens.

Brierly, A. S., Ward, P., Watkins, J. L., and Goss, C. (**1998**). ''Acoustic discrimination of Southern Ocean zooplankton,'' Deep Sea Res. II **45(7)**, 1155–1173.

Chu, D., Stanton, T. K., and Wiebe, P. H. (**1992**). ''Frequency dependence of sound backscattering from live individual zooplankton,'' ICES J. Mar. Sci., **49**, 97–106.

Demer, D. A. (**1994**). ''Accuracy and precision of acoustic surveys of Antarctic krill,'' Ph.D. thesis, UCSD.

Demer, D. A., and Martin, L. V. (**1995**). ''Zooplankton target strength: Volumetric or areal dependence?'' J. Acoust. Soc. Am. **98**, 1111–1118.

Devereaux, R. F. (**1953**). ''Isaacs-Kidd midwater trawl,'' Scripps Institution of Oceanography, Refer. 53–3.

Ehrenberg, J. E. (**1979**). ''A comparative analysis of *in situ* methods for directly measuring the acoustic target strength of individual fish,'' IEEE J. Ocean Eng. **OE-4**, 141–152.

Ehrenberg, J. E. (**1989**), ''A Review of Target Strength Estimation Techniques,'' in *Underwater Acoustic Data Processing* (Kluwer, Boston), pp. 161–175.

Fernandes, P. G., and Simmonds, E. J. (**1996**). ''Practical approaches to account for receiver delay and the TVG start time in the calibration of the Simrad EK500,'' ICES C. M. B:17, 8 pp (mimeo).

Foote, K. G. (**1983**). ''Maintaining precision calibrations with optimal copper spheres,'' J. Acoust. Soc. Am. **73**, 1054–1063.

Foote, K. G. (**1990**). ''Spheres for calibrating an eleven-frequency acoustic measurement system,'' ICES J. Mar. Sci. **46**, 284–286.

Foote, K. G. (**1991**). ''Summary of methods for determining fish target strength at ultrasonic frequencies,'' ICES J. Mar. Sci. **48**, 211–217.

Foote, K. G. (**1996**). ''Coincidence echo statistics,'' J. Acoust. Soc. Am. **99**, 266–271.

Greenlaw, C. F. (**1979**). ''Acoustical estimation of zooplankton populations,'' Limnol. Oceanogr. **24**, 226–242.

Greenlaw, C. F., and Johnson, R. K. (**1983**). ''Multiple-frequency Acoustical Estimation,'' Biol. Oceanogr. **2**, 227–252.

Hewitt, R. P., and Demer, D. A. (**1991**). ''Krill abundance,'' Nature (London), **353**, 310.

Hewitt, R. P., and Demer, D. A. (**1993**). ''Dispersion and abundance of Antarctic krill in the vicinity of Elephant Island in the 1992 austral summer,'' Mar. Ecol. Prog. Ser. **99**, 29–39.

Holliday, D. V. (**1977**). ''Extracting bio-physical information from the acoustic signatures of marine organisms,'' in *Oceanic Sound Scattering Prediction*, edited by N. R. Andersen and B. J. Zahuranec (Plenum, New York).

Kerr, D. E. (**1988**). *Propagation of Short Radio Waves* (McGraw-Hill, New York, 1951), reprinted (Peninsula, Los Altos, CA).

MacLennan, D. N. (**1987**). ''Time-varied-gain functions for pulsed sonars,'' J. Sound Vib. **110**, 511–522.

MacLennan, D. N., and Simmonds, E. J. (**1992**). *Fisheries Acoustics* (Chapman and Hill, London).

Madureira, L. S. P., Ward, P., and Atkinson, A. (**1993**). ''Differences in backscattering strength determined at 120 and 38 kHz for three species of Antarctic macroplankton,'' Mar. Ecol. Prog. Ser. **93**, 17–24.

Martin, L. V., Stanton, T. K., Wiebe, P. H., and Lynch, J. F. (**1996**). ''Acoustic classification of zooplankton,'' ICES J. Mar. Sci. **53**(2), 217–224.

Medwin, H., and Clay, C. S. (**1998**). *Fundamentals of Acoustical Oceanography* (Academic, San Diego).

Rice, J. A. (**1988**). *Mathematical Statistics and Data Analysis* (Duxbury Press, Belmont, CA).

Senturia, S. D. and Wedlock, B. D. (**1975**). *Electronic Circuits and Applications* (Wiley, New York).

Simmonds, E. J., Armstrong, F., and Copland, P. J. (**1996**). ''Species identification using wideband backscatter with neural network and discriminant analysis,'' ICES J. Mar. Sci. **53**(2), 189–196.

Simrad (**1996**). ''Simrad EK500 Scientific Echo Sounder Instruction Manual,'' Simrad Subsea A/S, Horten, Norway.
Soule, M. A., Barange, M., and Hampton, I. (**1995**) ''Evidence of bias in estimates of target strength obtained with a split-beam echo-sounder,'' ICES J. Mar. Sci. **52**, 139–144.
Soule, M. A., Hampton, I., and Barange, M. (**1996**). ''Potential improvements to current methods of recognizing single targets with a split-beam echo-sounder,'' ICES J. Mar. Sci. **53**, 237–243.
Soule, M. A., Barange, M., Solli, H., and Hampton, I. (**1997**). ''Performance of a new phase algorithm for discriminating between single and overlapping echoes in a split-beam echo-sounder,'' ICES J. Mar. Sci. **54**, 934–938.
Stanton, T. K., Wiebe, P. H., Chu, D., Benfield, M., Scanlon, L., Martin, L. V., and Eastwood, R. L. (**1994**). ''On acoustic estimates of zooplankton biomass,'' ICES J. Mar. Sci. **51**, 505–512.
Thompson, C. H., and Love, R. H. (**1996**). ''Determination of fish size distributions and areal densities using broadband low-frequency measurements,'' ICES J. Mar. Sci. **53**(2), 197–202.
Wilson, O. B. (**1988**). *Introduction to Theory and Design of Sonar Transducers* (Peninsula, Los Altos, CA).
Zakharia, M. E., Magand, F., Hetroit, F., and Diner, N. (**1996**). ''Wideband sounder for fish species identification at sea,'' ICES J. Mar. Sci. **53**(2), 203–208.

Holographic reconstruction of active sources and surface admittance in an enclosure

Young-Key Kim[a] and Yang-Hann Kim[b]
Center for Noise and Vibration Control, Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology, Science Town, Taejon 305-701, Korea

(Received 8 May 1998; revised 10 December 1998; accepted 11 December 1998)

Near-field acoustic holography (NAH) is applied to reconstruct an exterior sound field, as well as an interior sound field. A numerical method is employed to reconstruct the interior sound field of an arbitrary shaped enclosure. Reconstructed sound fields are usually used in identifying noise sources. However, until now, it has not been thoroughly verified whether the reconstructed sound field of an interior also shows source distributions. Interior sound fields have been reconstructed without any thought given to this question. In contrast to an exterior sound field, reflections are added to an interior sound field and therefore, the reconstructed sound field may misrepresent the source distribution. This paper addresses a way to solve this problem. A measurement method is proposed in order to distinguish between reflections and direct radiation from noise sources. © *1999 Acoustical Society of America.* [S0001-4966(99)05703-3]

PACS numbers: 43.60.Sx, 43.60.Pt, 43.58.Bh [ANN]

INTRODUCTION

To effectively control the interior noise of a vehicle or fuselage, it is essential to find the primary sources of the noise. In contrast to passive reactions, such as reflections and absorption, primary noise sources actively radiate sound into an enclosure. Therefore, if one eliminates primary sources, perfect control will have been obtained. The main objective of this paper is to identify these primary sources, which can be defined as *active* sources due to their activity.

Although near-field acoustic holography (NAH)[1] is well known for its source identification capability, it has limitations in estimating active sources in an enclosure. Historically, NAH has been applied to an exterior sound field, but a numerical approach extended its area to an interior sound field. Veronesi and Maynard[2] used a numerical method in order to model an arbitrarily shaped surface. In addition, Gardner and Bernhard[3] presented the same derivation by using the inverse of the Helmholtz integral equation for an enclosure. However, it is noteworthy that both of their purposes were to reconstruct a mixed sound field, rather than to distinguish between active sources and passive reactions. Therefore, their methods cannot be used to find the active sources of an enclosure.

In this paper, active sources are simply identified by eliminating passive reactions from a reconstructed sound field. The following sections describe this elimination procedure. Since the passive reactions are directly controlled by surface characteristics, such as surface admittance, the procedure includes a technique for the measurement of surface admittance.

There are five sections in this paper. Near-field acoustic holography is briefly reviewed in Sec. I. Theoretical derivations are described in Secs. II and III. Sections IV and V are for numerical simulations and summary.

[a]Electronic mail: youngkey@sam.kaist.ac.kr
[b]Electronic mail: yhkim@sorak.kaist.ac.kr

I. HOLOGRAPHIC RECONSTRUCTION OF A SOUND FIELD (Refs. 1–3)

Near-field acoustic holography (NAH) is derived from the Helmholtz integral equation. It is the inverse of this equation. For an arbitrarily shaped enclosure, NAH is numerically implemented by using the boundary element method (BEM).

The Helmholtz integral equation gives the sound-pressure field $p(\mathbf{r})$ at a field point $\mathbf{r}$ in terms of surface pressure $p(\mathbf{r}_s)$ and a normal component of the surface velocity $v_n(\mathbf{r}_s)$ on surface S (Fig. 1),

$$\gamma(\mathbf{r})p(\mathbf{r}) = -\int_S p(\mathbf{r}_s')\frac{\partial G}{\partial n}(\mathbf{r}-\mathbf{r}_s')dS + i\rho ck\int_S v_n(\mathbf{r}_s')G(\mathbf{r}-\mathbf{r}_s')dS, \quad (1)$$

where $G(\mathbf{r}-\mathbf{r}_s') = \exp(ikR)/4\pi R$, $R=|\mathbf{r}-\mathbf{r}_s'|$, and $\partial/\partial n$ is the normal derivative with respect to the surface. The speed of sound is c, the fluid mass density is ρ, and $k=\omega/c = 2\pi/\lambda$, with ω the frequency and λ the radiated wavelength. The coefficient $\gamma(\mathbf{r})$ represents the solid angle, which is 1 for a point in a volume V and 1/2 for the surface having a unique tangent plane. The exact value of this coefficient can be obtained by Cauchy's principal value.

Equation (1) can be numerically implemented by using the boundary element method. In this method, the geometric and acoustical quantities are interpolated by using shape functions and the quantities at nodes (for details, see a textbook such as that in Ref. 4). If we summarize the results, they have the form,

$$[D]_s\{p\}_s = [M]_s\{v_n\}_s, \quad (2)$$

 0001-4966/99/105(4)/2377/7/$15.00 © 1999 Acoustical Society of America

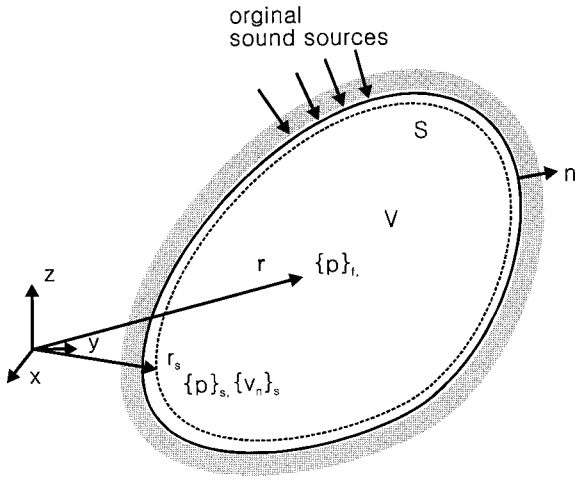


FIG. 1. The definition of geometric conditions and acoustic variables. The vectors $\{p\}_s$ and $\{v_n\}_s$ denote surface pressure and velocity on the surface while $\{p\}_f$ denotes field pressure. Vector **n** indicates a normal direction to the surface.

$$\{p\}_f = -[D]_f\{p\}_s + [M]_f\{v_n\}_s, \tag{3}$$

where $\{p\}_s$ and $\{v_n\}_s$ denote the surface pressure and surface normal velocity vectors, and $\{p\}_f$ is the field pressure vector. $[D]_s$ and $[M]_s$ are the transfer matrices, which are called dipole and monopole matrices. $[D]_f$ and $[M]_f$ are dipole and monopole matrices for field pressures. If one knows the boundary conditions of a sound field, such as surface pressure and surface velocity, or a relation between these two, then one can calculate the field pressure $\{p\}_f$ from Eqs. (2) and (3).

Finally, the basic equations of NAH are derived from Eqs. (2) and (3). If $\{p\}_s$ and $\{v_n\}_s$ are eliminated one after the other, we have

$$\{p\}_f = ([M]_f - [D]_f[D]_s^{-1}[M]_s)\{v_n\}_s \equiv [T]_{v_n}\{v_n\}_s, \tag{4}$$

$$\{p\}_f = (-[D]_f + [M]_f[M]_s^{-1}[D]_s)\{p\}_s \equiv [T]_p\{p\}_s, \tag{5}$$

provided that $[D]_s^{-1}$ and $[M]_s^{-1}$ exist.

Equations (4) and (5) are direct relations between field pressure $\{p\}_f$, and surface acoustic variables $\{p\}_s$ and $\{v_n\}_s$. Therefore, if we measure field pressure, then we can calculate the surface pressure and the surface velocity. In addition, field pressures at other field points can be calculated by using Eq. (3).

One thing that should be noted is that singular frequencies are in the NAH. The matrices $[D]_s$ and $[M]_s$ become singular at the characteristic frequencies of an enclosure.[2,5] Theoretically, these singularities only occur at the characteristic frequencies. However, practically, the measurement errors are amplified around these frequencies. In order to reduce these errors, the singular value decomposition is usually accompanied by the NAH.[2,5] It eliminates less contributing modes in inverse processes and reduces the amplification of errors. However, we will not go into detail on this topic since our focus is on proposing a method using NAH, rather than improving it.

II. HOLOGRAPHIC RECONSTRUCTION OF ACTIVE SOURCES

A. Active sources and passive reactions

Although NAH simply reconstructs a whole sound field, it does not distinguish between active sources and passive reactions. Mathematically, the reason can be found in the fact that NAH does not use a complete boundary condition in representing a sound field. This fact is not directly mentioned in the theoretical derivations, but Eqs. (4) and (5) are based on simple Neumann and Dirichlet conditions.[6] These equations represent the sound field due to velocity and pressure boundary conditions. Therefore, NAH is not concerned with the physical boundary conditions, which correctly describe how much the boundaries radiate and reflect sound. As a result, the NAH simply estimates equivalent Neumann and Dirichlet conditions in order to reconstruct a sound field by using them.

Generally, two kinds of conditions are needed to represent a real surface: active and passive. In other words, most surfaces actively radiate sound waves, as well as passively reacting to incident waves. A good example is the vibrating panel covered with absorbent material, which partially radiates and absorbs sound. Although the sound field due to this panel can be represented with equivalent boundary conditions, such as surface velocity or pressure, they do not show how much it radiates and absorbs sound. Therefore, both active and passive terms should be included in a boundary condition in order to represent a real surface.

In this paper, we represented a boundary with the Cauchy condition,[6,7]

$$\alpha(\mathbf{r}_s)p(\mathbf{r}_s) + \beta(\mathbf{r}_s)v_n(\mathbf{r}_s) = f(\mathbf{r}_s), \tag{6}$$

where $\alpha(\mathbf{r}_s)$ and $\beta(\mathbf{r}_s)$ are complex numbers.

It is rearranged in order to conveniently impose physical meaning. If one divides both sides of Eq. (6) by $\beta(\mathbf{r}_s)$, then one has

$$v_n(\mathbf{r}_s) = A(\mathbf{r}_s)p(\mathbf{r}_s) + S(\mathbf{r}_s), \tag{7}$$

with $S(\mathbf{r}_s) \equiv f(\mathbf{r}_s)/\beta(\mathbf{r}_s)$ and $A(\mathbf{r}_s) \equiv -\alpha(\mathbf{r}_s)/\beta(\mathbf{r}_s)$, provided that $\beta(\mathbf{r}_s)$ is not zero.

In Eq. (7), it is easy to notice that the first term represents passive reactions from a locally reacting surface.[8] Surface admittance $A(\mathbf{r}_s)$ determines the ratio between imposed surface pressure and its response. Therefore, if one uses Eq. (7) in order to model a locally reacting surface, then the first term will represent passive components of a surface velocity. The second term is the remainder. Since the passive component is excluded by the first term, the second term can be regarded as an active source.

Equation (7) can also be generalized for an extensively reacting surface. In this case, the Cauchy condition can be written as

$$v_n(\mathbf{r}_s) = \int_S A(\mathbf{r}_s|\mathbf{r}_s')p(\mathbf{r}_s')dS + S(\mathbf{r}_s), \tag{8}$$

where the integral and the kernel $A(\mathbf{r}_s|\mathbf{r}_s')$ represent the passive reactions of a whole surface. Physically, the kernel

$A(\mathbf{r}_s|\mathbf{r}'_s)$ is the surface admittance of an extensively reacting surface.

B. Estimation of active sources

The sound field, due to the Cauchy condition, can be derived by the substitution of the Cauchy condition [Eq. (8)] into Eq. (1),

$$\gamma(\mathbf{r})p(\mathbf{r}) = -\int_S \left\{ \frac{\partial G}{\partial n}(\mathbf{r}-\mathbf{r}'_s) - i\rho ck \int_S A(\mathbf{r}''_s|\mathbf{r}'_s) \times G(\mathbf{r}-\mathbf{r}''_s)dS \right\} p(\mathbf{r}'_s)dS + i\rho ck \int_S S(\mathbf{r}'_s) \times G(\mathbf{r}-\mathbf{r}'_s)dS. \quad (9)$$

If one rewrites Eq. (9) in a matrix equation using the same matrices in Eqs. (2) and (3), one has

$$\{p\}_f = (-[D]_f + [M]_f[A])\{p\}_s + [M]_f\{S\}_s, \quad (10)$$

where vector $\{S\}_s$ is the source strength vector, and $[A]$ is the admittance matrix. If one uses a locally reacting condition [Eq. (7)] instead of an extensively reacting one, $[A]$ will be a diagonal matrix.

Furthermore, if one eliminates the surface pressure $\{p\}_s$ in Eq. (10) by using Eq. (5), one has the basic equation of the proposed method,

$$\{S\}_s = [M]_f^+[[D]_f - [M]_f[A] + [T]_p][T]_p^+\{p\}_f, \quad (11)$$

where $^+$ indicates the pseudo inverse, which is defined as $[G]^+ = ([G]^H[G])^{-1}[G]^H$. The superscript H denotes the Hermitian operation.

Equation (11) relates active source strength, $\{S\}_s$, and field pressure, $\{p\}_f$. The transfer matrix between them is rather complex, but the only unknown variable is matrix $[A]$. It shows that we can estimate active sources from the measured sound field, if the surface admittance is available.

III. A TECHNIQUE FOR THE ESTIMATION OF SURFACE ADMITTANCE

As shown in Eq. (11), one needs to measure surface admittance in order to find active sources. The following section describes a technique that estimates surface admittance of a whole surface. In this paper, we focus on a locally reacting surface because of its simplicity and practical importance. However, we will also show how it can be generalized.

A locally reacting assumption limits the application area of our method. For example, a flexible panel is not locally reacting since the whole panel responds to point excitation. However, practically, this assumption is acceptable in many cases. Most absorbent materials are assumed to be locally reacting.[8] Furthermore, most of the response from a flexible panel is negligible, except at resonance frequencies.

Surface admittance has been measured by several methods, such as the two-sensor method[9] and the spatial Fourier transform method.[10] However, they were for small or uniform specimens. It was impossible to use them for the measurement of a whole surface. For example, if one measures the admittance of an entire surface using the two-sensor method, then one will need hundreds of specimens. In addition, careful attention must be paid when taking specimens. Therefore, we developed a measurement technique for an entire surface.

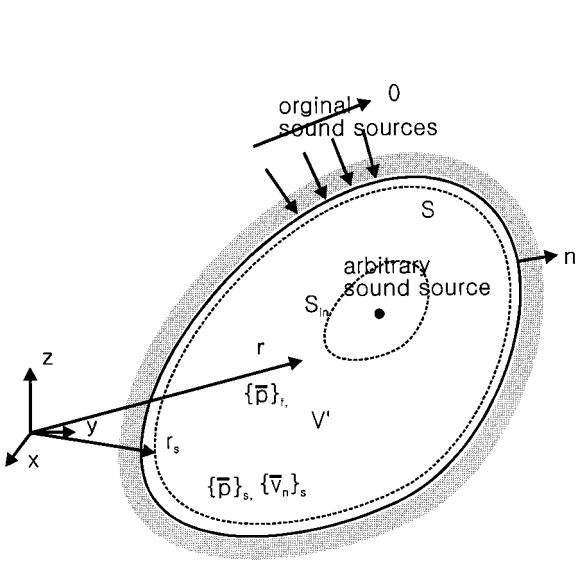


FIG. 2. The measurement configuration for the admittance measurement method. The inner surface, S_{in}, and the outer surface, S, construct volume V'. In order to measure surface admittance, near-field acoustic holography reconstructs sound velocity and pressure on surface S.

The proposed technique is similar to the two-sensor method; however, we generalized it for three-dimensional space. The measurement steps are simple. First, a surface is excited with an arbitrary sound source after active sources are turned off. For example, the engine of a car is turned off and the enclosure is excited with a loud speaker. Then NAH is applied in this enclosure in order to estimate surface velocity and pressure. For a locally reacting surface, the ratio of the reconstructed surface velocity and pressure gives the surface admittance.

Figure 2 shows the conceptual diagram of the admittance measurement technique. In this configuration, an arbitrary source excites an enclosure, and NAH is applied to the volume between surfaces S and S_{in}.

If we turn off active sources on boundary S, the Cauchy condition of Eq. (7) becomes

$$\overline{v_n}(\mathbf{r}_s) = A(\mathbf{r}_s)\overline{p}(\mathbf{r}_s), \quad (12)$$

where $\overline{v_n}(\mathbf{r}_s)$ and $\overline{p}(\mathbf{r}_s)$ represent the surface velocity and pressure which are excited by an arbitrary sound source (Fig. 2).

Furthermore, Eq. (12) can be expressed as a matrix equation,

$$\{\overline{v_n}\}_s = [A]\{\overline{p}\}_s, \quad (13)$$

where the admittance matrix $[A]$ is diagonal. Therefore, the ratio of surface velocity, $\{\overline{v_n}\}_s$ and the surface pressure, $\{\overline{p}\}_s$ gives the diagonal elements of the admittance matrix $[A]$.

For an extensively reacting surface, Eq. (8) can be rewritten as

$$\overline{v_n}(\mathbf{r}_s) = \int_S A(\mathbf{r}_s|\mathbf{r}'_s)\overline{p}(\mathbf{r}'_s)dS, \quad (14)$$

and the matrix equation is

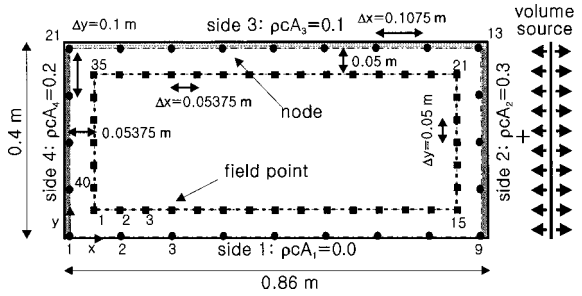


FIG. 3. A rectangle with a vibrating source on side 2. A total of 24 nodes are on the boundaries. The source strength is inversely calculated from the field pressures, which are taken at 40 field points.

$$\{\overline{v_n}\}_s=[A']\{\overline{p}\}_s, \tag{15}$$

where $[A']$ is a nondiagonal matrix; $A(\mathbf{r}_s|\mathbf{r}_s')\neq 0$ provided $\mathbf{r}_s\neq\mathbf{r}_s'$.

Equation (15) is similar to Eq. (13), but there are $N\times N$ unknowns since $[A']$ is not diagonal. Therefore, it is impossible to calculate these unknowns from the simple ratio between surface velocity, $\{\overline{v_n}\}_s$, and surface pressure, $\{\overline{p}\}_s$.

Theoretically, it is possible to create more equations in order to estimate matrix $[A']$. For example, if we move an arbitrary exciting source from one point to another, then we will have two sets of equations:

$$\{\overline{v_n}\}_s^{(1)}=[A']\{\overline{p}\}_s^{(1)}, \tag{16}$$

$$\{\overline{v_n}\}_s^{(2)}=[A']\{\overline{p}\}_s^{(2)}, \tag{17}$$

where the superscripts denote the positions of the exciting sources. Furthermore, following the same procedure, it is possible to create N sets of equations. However, this procedure will take a very long time since we would have to apply NAH to each sound field.

IV. NUMERICAL SIMULATIONS FOR A TWO-DIMENSIONAL SOUND FIELD IN A RECTANGLE

A. Two-dimensional sound fields

Two-dimensional sound fields were chosen to verify the proposed method. Although a two-dimensional sound field is not common in practice, it is simple and easy to simulate. It requires a smaller number of nodes in numerical modeling than a three-dimensional sound field does.

For a two-dimensional sound field, Green's function in Eq. (1) is different,

$$G(\mathbf{r}-\mathbf{r}_s')=\{iJ_0(kR)-Y_0(kR)\}/4, \tag{18}$$

where $J_0(kR)$ and $Y_0(kR)$ are the Bessel functions of the first and second kinds of order zero. The singularity of $Y_0(kR)$ can be sufficiently reduced if we expand it into a logarithmic function and remainders.[11] For numerical formulations, linear shape functions were used.

Sound fields in a rectangle (Fig. 3) were used in these simulations. The admittance of each wall was assumed to be purely real, $\rho cA_1=0.0$, $\rho cA_2=0.3$, $\rho cA_3=0.1$ and $\rho cA_4=0.2$. A sound source was located on wall 2. For simplicity, we assumed that the source was in-phase along wall 2 (Fig. 3). This source can be interpreted as a vibrating wall covered with absorbent material. A total of 24 nodes were selected among the boundaries. The highest frequency, which is allowed for this node separation, was 797 Hz. The resonance frequencies of the rectangle are listed in Table I.

TABLE I. Resonance frequencies of a rectangle with rigid boundaries. The length and width of the rectangle are 0.86 m and 0.4 m, respectively.

No.	Mode number (m, n)	Resonance frequency (Hz)
1	(1,0)	199.4
2	(2,0)	398.8
3	(0,1)	428.8
4	(1,1)	472.9
5	(2,1)	582.6
6	(3,0)	598.3
7	(3,1)	736.0

In order to simulate the sound fields, the modal summation technique[8] was used. Since the admittance of the boundaries is not very small, we solved full modal coupled equations. A total of 1800 (60×30) modes were used in this calculation.

B. The estimation of source strength with known admittance

First, source strength was estimated with known admittance. In fact, it is the second step of the proposed method. We verified this step separately since it includes the main idea of the proposed method. The first step and the total procedure are verified in Secs. IV C–D. Figure 3 shows the estimation conditions. Twenty-four nodes are on the boundary. In addition, there are 40 points where we took field pressure. The number of field points is 1.67 times that of the surface nodes.

Figure 4 shows the example of estimated source strength (solid line) at 400 Hz. This graph was plotted from side 1 to side 4. Furthermore, the estimated source strength was compared with the reconstructed surface velocity (thick, dashed line). This plot compares the performance of the proposed method to that of NAH.

In Fig. 4, estimated source strength largely agrees with true source strength and shows that there is a vibrating source on side 2. However, surface velocity, which was reconstructed by using NAH, fails to indicate source location. It also has nonzero magnitudes on sides 3 and 4 due to pas-

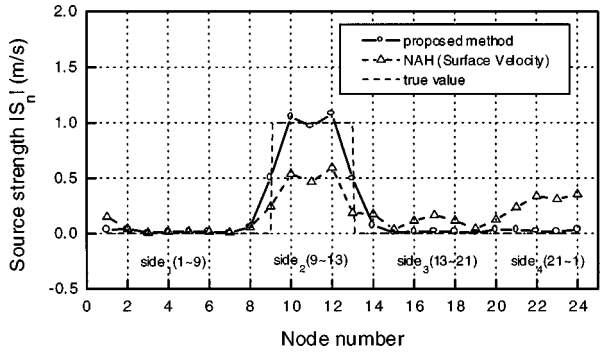
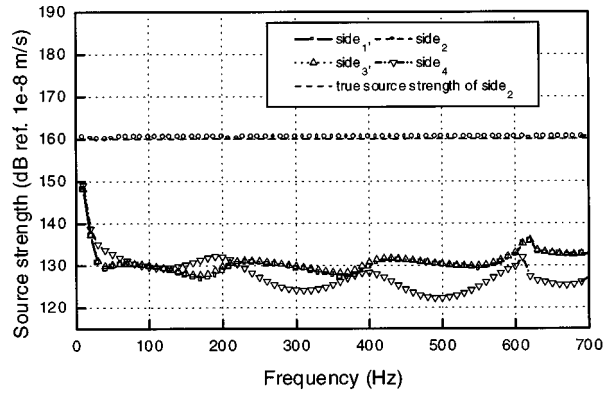
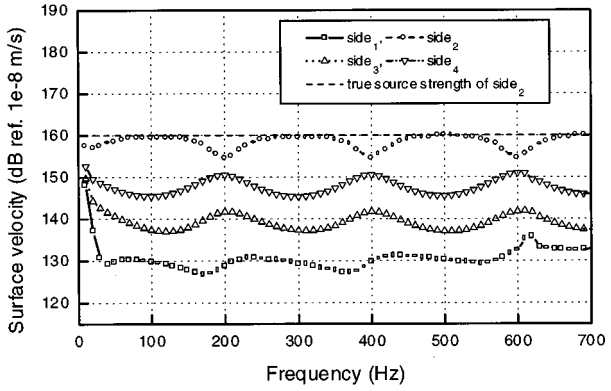


FIG. 4. Estimated source strength and surface velocity at 400 Hz. This graph compares the performance of the proposed method to that of near-field acoustic holography (NAH).



(a)



(b)

FIG. 5. (a) The average source strength reconstructed by the proposed method; (b) the average surface velocity reconstructed by near-field acoustic holography (NAH).

sive reactions of the boundary. In addition, estimated source strength was smoothed at nodes 9 and 13 since we used a finite number of nodes.

Figure 5 shows the frequency characteristics of estimated source strength and surface velocity. For each frequency, we averaged source strength and surface velocity on each side. This graph shows them on a decibel scale with the reference 10^{-8} m/s. In Fig. 5(a), the average source strength of side 2 is about 30 dB higher than that of the other sides. This locates the sound source on side 2 rather well. The small error around 620 Hz occurred due to practical limitations of the measurement system. At this frequency, a singular sound field, which has zero sound pressure along the line made by field sensors, can occur. It can be easily checked by the characteristic frequency of the rectangle made by field sensors (Fig. 3).

Figure 5(b) is the average surface velocity which was estimated by NAH. It shows that NAH fails to indicate source locations. The average surface velocity has large magnitudes on sides 3 and 4 where there is no source. In addition, it varies along the frequency axis. It occurred as a result of passive reactions, which depend on surface admittance and the shape of the enclosure. Since sound source is in-phase in the y-direction, the frequency characteristics of the enclosure depend on the x-axis mode; for example, $(m,0)$, $m=1,2\cdots$ modes. Therefore, the variation is periodic

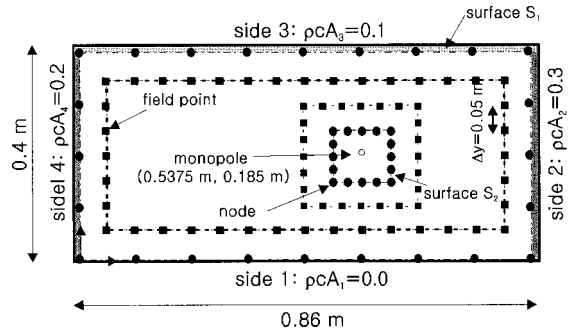
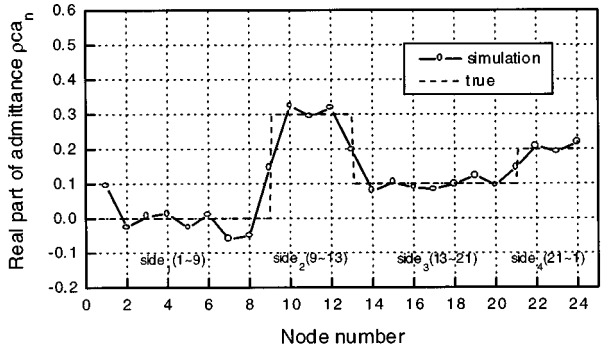


FIG. 6. A two-dimensional rectangle with a monopole source in it. A total of 40 nodes are on surface S_1 and the imaginary surface S_2. Surface velocity and pressure are inversely calculated from the field pressures, which are taken at 64 field points.

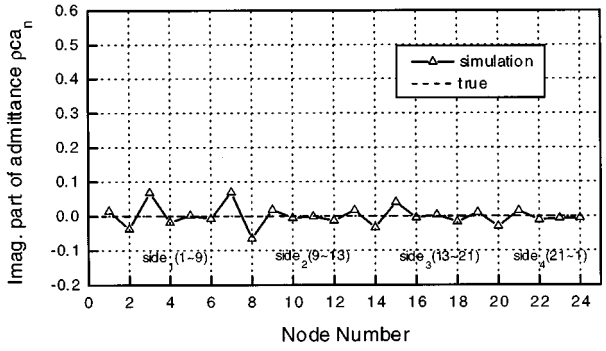
with the period of 200 Hz, which is the first resonance frequency of the rectangle.

C. Measurements of wall admittance

Next, we verified the admittance measurement method. Figure 6 shows an estimation condition. There were 24 nodes on boundary S_1, and the source on side 2 was turned off in this case. An additional 16 nodes were placed on an imaginary surface S_2. The field pressures were taken at 64 measurement points. The number of measurement points is 1.6 times that of the surface nodes. In order to excite the



(a)



(b)

FIG. 7. Estimated wall admittance at 400 Hz. (a) Real part; (b) imaginary part.

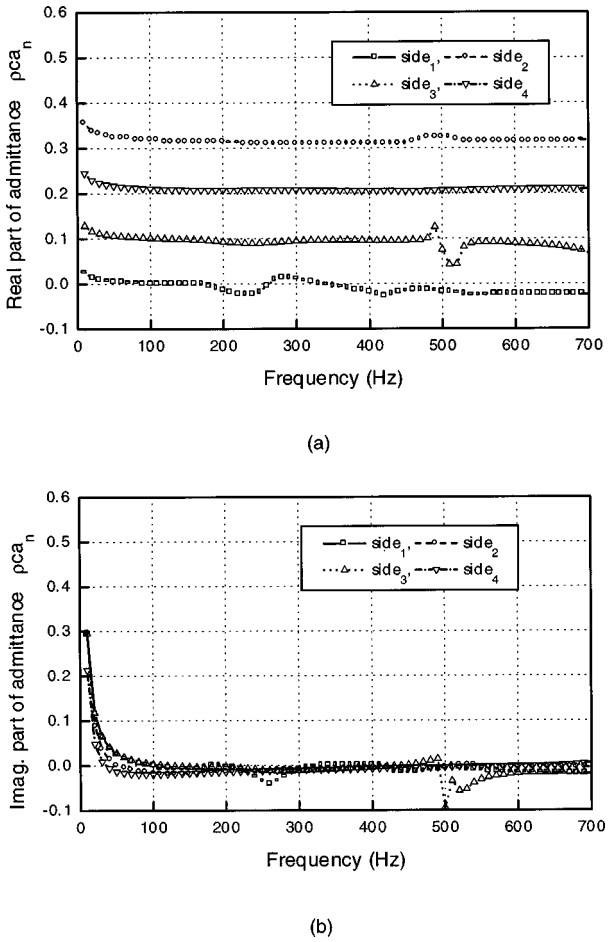


FIG. 8. The average admittance of each wall. (a) Real part; (b) imaginary part. The true values are $\rho ca_1 = 0.0 + 0.0i$, $\rho ca_2 = 0.3 + 0.0i$, $\rho ca_3 = 0.1 + 0.0i$, $\rho ca_4 = 0.2 + 0.0i$.

enclosure, a monopole source was located at (0.5375 m, 0.185 m). The modal summation technique was employed in calculating a sound field. A total of 1800 (60×30) modes were summed up.

Figure 7 shows an example of the estimated admittance at 400 Hz. The estimated admittance shows a high level of agreement with the true admittance at most nodes. The average admittance of each wall is shown in Fig. 8. At most frequencies, it largely agrees with the true admittance, but shows rather large errors at 265 and 500 Hz. If we calculate wavelengths multiplied by $(n+1/2)$ at these frequencies, then we know that they are the multiple of the distances between sides 2 and 4. This means that input impedance at these frequencies is very high. Therefore, a strong near field is formed around the excitation source and causes large errors in the modal summation. If we increase the number of modes, the calculated field will converge into a true sound field.

D. Estimation of sources with measured admittance

Finally, we verified the total procedure of the proposed method. Figure 9 shows estimated source strength. It was estimated using the estimated admittance of Fig. 6. It also largely agrees with true source strength, but it shows small errors at 0 and 500 Hz. They are the effect of admittance errors. If one compares Fig. 9 to Fig. 8, one will see that both have errors at the same frequencies. However, they were not largely amplified. Estimated source strength shows the source locations correctly.

V. SUMMARY

This paper proposed a measurement method which estimated active sources on boundaries. First, the Cauchy condition was used to model active sources and surface admittance. Then this condition was estimated by the proposed method. In addition, the admittance measurement technique was proposed. We showed that it estimates the admittance of any surface as far as it is locally reacting. The simple numerical simulations for a two-dimensional sound field were performed to verify the proposed method. Estimated admittance and source strengths largely agreed with true values. Furthermore, the proposed method was compared to near-

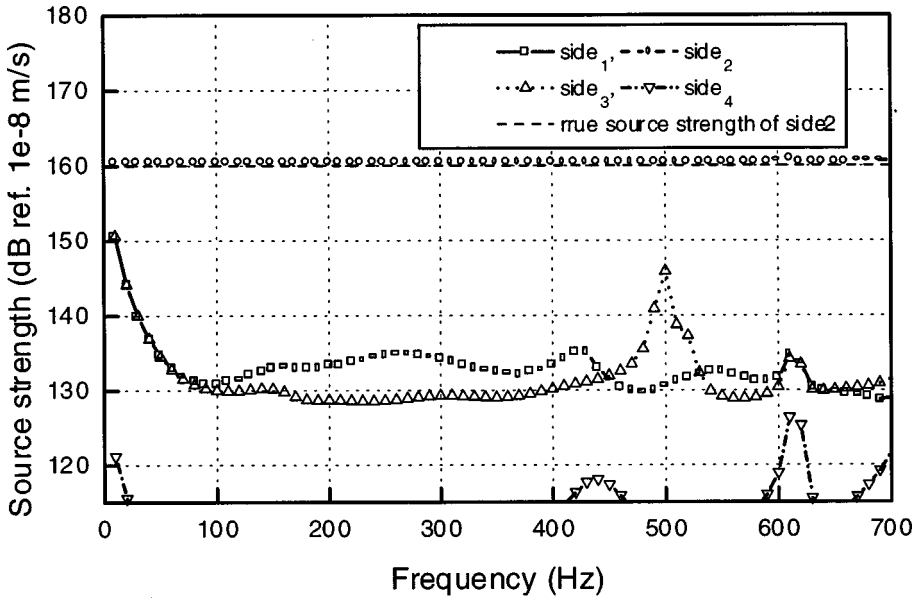


FIG. 9. The average source strength reconstructed by the proposed method. Average source strength indicates that there is a source on side 2. In this case, the estimated admittance (Fig. 8) is used in reconstruction.

field acoustic holography. We showed that only our method was able to estimate source locations correctly.

[1] J. D. Maynard, E. G. Williams, and Y. Lee, ''Near field acoustic holography: I. Theory of generalized holography and the development of NAH,'' J. Acoust. Soc. Am. **78**, 1397–1413 (1985).

[2] W. A. Veronesi and J. D. Maynard, ''Digital holographic reconstruction of sources with arbitrarily shaped surfaces,'' J. Acoust. Soc. Am. **85**, 588–598 (1989).

[3] B. K. Gardner and R. J. Bernhard, ''A noise source identification technique using an inverse Helmholtz integral equation method,'' Trans. ASME, J. Vib. Acoust. Stress **110**, 84–90 (1988).

[4] R. D. Ciskowski and C. A. Brebbia, *Boundary Element Methods in Acoustics* (Elsevier Applied Science, London, 1991), pp. 13–76.

[5] 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**, 3003–3016 (1996).

[6] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics, Part I* (McGraw-Hill, New York, 1968), pp. 676–679.

[7] L. L. Beranek and I. L. Ver, *Noise and Vibration Control Engineering* (Wiley, New York, 1992), pp. 149–150.

[8] A. D. Pierce, *Acoustics* (McGraw-Hill, New York, 1981), pp. 110–111, 288–291.

[9] A. F. Seybert, ''Two-sensor method for the measurements of sound intensity and acoustic properties in duct,'' J. Acoust. Soc. Am. **83**, 2233–2239 (1988).

[10] M. Tamura, ''Spatial Fourier transform method of measuring reflection coefficients at oblique incidence. I: Theory and numerical examples,'' J. Acoust. Soc. Am. **88**, 2259–2264 (1990).

[11] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965), p. 364.

Frequency glides in the impulse responses of auditory-nerve fibers

Laurel H. Carney,[a] Megean J. McDuffy, and Ilya Shekhter
Boston University Hearing Research Center and Department of Biomedical Engineering, Boston University, 44 Cummington Street, Boston, Massachusetts 02215

(Received 24 July 1998; revised 2 November 1998; accepted 15 January 1999)

Previous reports of frequency modulations, or glides, in the impulse responses of the auditory periphery have been limited to analyses of basilar-membrane measurements and responses of auditory-nerve (AN) fibers with best frequencies (BFs) greater than 1.7 kHz. These glides increased in frequency as a function of time. In this study, the instantaneous frequency as a function of time was measured for impulse responses of AN fibers in the cat with a range of BFs (250–4500 Hz). Impulse responses were estimated from responses to wideband noise using the reverse-correlation technique. The impulse responses had increasing frequency glides for fibers with BFs greater than 1500 Hz, nearly constant frequency as a function of time for BFs between 750 and 1500 Hz, and decreasing frequency glides for BFs below 750 Hz. Over the levels tested, the glides for fibers at all BFs were nearly independent of stimulus level, consistent with previous reports of impulse responses of the basilar membrane and AN fibers. Implications of the different glide directions observed for different BFs are discussed, specifically in relation to models for the auditory periphery as well as for the derivation of impulse responses for the human auditory periphery based on psychophysical measurements. © *1999 Acoustical Society of America.* [S0001-4966(99)03204-X]

PACS numbers: 43.64.Bt, 43.64.Pg, 43.66.Ba [RDF]

INTRODUCTION

The properties of auditory nerve (AN) fibers can be characterized using their responses to impulses, or clicks, as well as to wideband noise. These properties have implications for our understanding of signal processing in the auditory periphery and of the neural cues provided by the periphery to the central nervous system. Responses of neurons to wideband noise stimuli have been studied for many years in an effort to characterize both linear and nonlinear response properties (e.g., deBoer, 1967; deBoer and deJongh, 1978; Marmarelis and Marmarelis, 1978; Eggermont *et al.*, 1983; Carney and Yin, 1988; Eggermont, 1993). The cross-correlation function of the response of the system to a wideband noise with the stimulus waveform provides an estimate for the system's first-order Wiener kernel (Lee and Schetzen, 1965). For a linear system, the first-order kernel is also the impulse response.

This cross-correlation analysis technique has been successfully applied to mechanical measurements of basilar-membrane motion (deBoer and Nuttall, 1997; Recio *et al.*, 1997). These measurements have characterized the impulse responses as narrow-band signals having a frequency modulation, or glide. For these mid- to high-frequency ($>\sim$1.7 kHz) fibers or places, the carrier frequency of the tuned impulse response increases as a function of time from a frequency lower than the best frequency (BF, the frequency to which the system is most responsive[1]) up to the BF. The glides estimated from responses to wideband noise are consistent with impulse responses measured in the same cochleae using click stimuli (deBoer and Nuttall, 1997). The increasing frequency glides are also consistent with other reports of mechanical measurements of click responses from the base of the cochlea (Robles *et al.*, 1976; Recio *et al.*, 1998). deBoer and Nuttall (1997) referred to impulse responses based on click responses as *direct* impulse responses, and those estimated from noise responses as *indirect* impulse responses.

Characterization of the tuning of AN fibers based on responses to wideband noise requires an extension of the cross-correlation technique. Because the response of a fiber is a series of action potentials, rather than a continuous-time waveform, one cannot perform a simple cross correlation between the stimulus waveform and the response. However, a spike-triggered average of the stimulus waveform preceding each action potential yields a derived impulse response, the reverse-correlation (revcor) function (deBoer, 1967; deBoer and Kuyper, 1968; Møller, 1977; deBoer and deJongh, 1978; Marmarelis and Marmarelis, 1978; Eggermont *et al.*, 1983). Equivalently, a cross correlation between the noise stimulus waveform and the peri-stimulus time (PST) histogram constructed from responses to multiple repetitions of the stimulus yields an estimate of the impulse response (Møller, 1977).

The reverse-correlation procedure is only successful for relatively low-frequency AN fibers (BFs $<\sim$4000 Hz) because it requires phase-locking of the fiber to the fine structure of the stimulus waveform. Revcor functions for low-frequency AN fibers are consistent with click responses measured for the same fibers (Carney and Yin, 1988), indicating that this technique for the indirect estimate of the AN impulse response is consistent with direct estimates of impulse responses.

[a]Author to whom correspondence should be addressed. Electronic mail: carney@bu.edu

Pseudolinear characterizations of AN fibers across a range of stimulus levels provide a characterization of AN fibers that is adequate for predicting temporal response properties to resonant stimuli at different sound levels (Carney and Yin, 1988). A population of revcor functions has served as a basis for a pseudolinear model for a population of AN fibers (Carney and Yin, 1988) as well as for a nonlinear, time-varying model for AN responses to arbitrary stimuli (Carney, 1993). However, these models were based on simple parameterizations of the revcor functions as gammatone functions. The gammatone models used to characterize AN impulse responses have constant frequency as a function of time, and thus do not accurately represent impulse responses that contain frequency glides.

Reports of frequency glides in the revcor functions of AN fibers have been consistent with the descriptions of the direct and indirect impulse responses of the basilar membrane. Møller and Nilsson (1979; also Møller, 1981, 1983) illustrated an increasing frequency glide in the revcor functions for two rat AN fibers with BFs of approximately 2 and 3 kHz. deBoer and Nuttall (1997) reported increasing frequency glides in revcor functions for AN fibers with BFs as low as 3 kHz. They also reported increasing frequency glides in impulse responses derived from the phase measurements of tuning curves of AN fibers with BFs as low as 1.76 kHz (deBoer and Nuttall, 1997).

The goal of this study was to characterize the impulse responses of AN fibers with low to moderate BFs, ranging from approximately 250 to 4500 Hz. Across this frequency range, it is possible to measure indirect impulse responses using the reverse-correlation technique. Revcor functions of a large population of AN fibers (Carney and Yin, 1988; Carney, 1990) were examined in this study. The nature of the frequency glides in these derived impulse responses and the variation of the glides as a function of stimulus level were investigated. The frequency glides are consistent with descriptions of the shift in peak frequency of tuning as a function of stimulus level (deBoer and Nuttall, 1997; Recio *et al.*, 1997); that is, changes in the latency of the envelope of the impulse response as a function of level, along with the underlying (but level-independent) frequency glide, result in apparent shifts in the tuning as a function of level. For example, for BFs with glides that increase as a function of time, responses to high-level stimuli are determined by impulse responses that peak during the early, low-frequency portion of the frequency glide. Alternatively, responses to low-level stimuli are determined by impulse responses that peak at longer latencies, during the high-frequency part of the glide. Thus, the increasing frequency glides at high BFs are consistent with a downward shift in the peak frequency of tuning as stimulus level increases. Because some low-BF fibers have been described as having shifts in tuning that go in the opposite direction, it was not anticipated that the glides would necessarily be similar to those described for high-BF measurements. For example, very low-BF (<1000 Hz) fibers have been described as having peak frequencies that shift *upwards* as stimulus level increases (Evans, 1981), which is opposite in direction to the shift seen for high-BF fibers (Evans, 1981; Møller, 1977) and basilar-membrane measurements made at high-BF places (Rhode, 1971; Sellick *et al.*, 1982; Robles *et al.*, 1986; deBoer and Nuttall, 1997; Recio *et al.*, 1997, 1998). The frequency modulations, or glides, in the impulse responses of AN fibers with BFs less than 1500 Hz reported in this study have different slopes than those previously reported for responses of fibers with higher BFs (Møller and Nilsson, 1979; Møller, 1981, 1983) or high-frequency places on the basilar membrane (deBoer and Nuttall, 1997; Recio *et al.*, 1997, 1998). Flat or decreasing frequency glides for fibers with BFs below 1500 Hz were observed in this study.

These frequency glides have interesting implications both for our understanding of signal processing in the auditory periphery and for cochlear models (Møller and Nilsson, 1979; deBoer and Nuttall, 1997; Recio *et al.*, 1998). The association of level-dependent shifts in peak frequency with the compressive nonlinearity make the time and level dependence of the frequency glides an important issue. As pointed out by deBoer and Nuttall (1997), the presence of the glide in both the indirect impulse response based on steady-state noise responses and in the actual impulse response indicates that it is not a reflection of a time-varying nonlinearity. In addition, they report that the glide is remarkably consistent over a wide range of stimulus levels, consistent with the reports of Recio *et al.* (1997, 1998). This was also the case for AN fibers in this study, regardless of the direction or slope of the glide. The frequency glide does not appear to be directly associated with the active process or the associated compressive nonlinearity; the frequency glide persists nearly unchanged across the wide range of levels tested (20–80 dB SPL), while the bandwidth and gain of the system vary.

Impulse responses with frequency glides have recently arisen in another area of auditory research: impulse responses with upward frequency glides (gammachirps) have been proposed to model the filters associated with asymmetrical psychophysical tuning curves (Irino and Patterson, 1997). However, because the asymmetry of psychophysical tuning curves varies with stimulus level (e.g., Rosen and Baker, 1994), Irino and Patterson (1997) proposed that the peripheral filters should have glides that vary in slope as a function of stimulus level. Such filters would be at odds with measurements of both direct and indirect impulse responses from the auditory periphery. The variations in asymmetry of psychophysical tuning curves as a function of level must arise from properties other than changes in the slope of frequency glides in the impulse responses.

I. METHODS

Noise responses of AN fibers were obtained from a data archive collected as part of previous studies (Carney and Yin, 1988; Carney, 1990). The physiological procedures used to obtain the functions are detailed in those papers. Briefly, the responses of AN fibers to wideband (10-kHz bandwidth) noises were recorded in barbiturate-anesthetized cats. Typically, 40 repetitions at each sound level of an 800-ms duration noise sample were obtained at 1-s intervals. Sound levels reported here are in dB SPL (computed over the 10-kHz bandwidth of the noise) for consistency with

other studies, rather than the peak-equivalent levels reported previously (Carney and Yin, 1988; Carney, 1990).

The results presented were based on revcor functions, typically computed for noise responses measured at two or more sound levels, for 214 fibers from 13 cats. The original (unsmoothed) revcor functions are accessible for further study via an ftp site (Appendix A).

Each revcor function was computed using the cross-correlation function for the stimulus waveform and the PST histogram of the noise response (deBoer and deJohngh, 1978). PST histograms were binned at 50-μs resolution, which matched the sampling time of the stimulus waveform. A one-octave bandpass filter, centered at the peak frequency of the spectrum of the revcor function, was used to smooth the revcor functions. The 8th-order bandpass filter was implemented in MATLAB™, and the output signal was compensated for the delay of the filter. The filtered revcor functions were tested to ascertain that the filter did not introduce any phase distortion that could have affected the positions of zero crossings; this test was accomplished by superimposing the filtered and original signals and carefully comparing the locations of the zero crossings.

At very low stimulus levels, the BF of the revcor function matches the characteristic frequency (CF) based on a threshold-tuning curve; however, as stimulus level increases, these two frequencies may diverge, depending on CF (e.g., Evans, 1981). The BF was estimated using the peak frequency of the spectrum of the revcor function, computed with the fast Fourier transform. When glides were studied across several SPLs, the average BF across the SPLs was used as the fiber's BF [Fig. 7(A)].

The instantaneous frequency of the revcor functions was estimated in two ways: using a Hilbert-transform technique and using zero crossings. Using Hilbert-transform techniques, the instantaneous frequency (IF) and envelope of an analytic signal could be determined (Appendix B). This technique was used to analyze the impulse responses derived from basilar-membrane responses to noise (deBoer and Nuttall, 1997; Recio *et al.*, 1997). A preliminary study found that this technique was also appropriate for AN revcor functions (Shekhter, 1997; Shekhter and Carney, 1997).

A limitation of the Hilbert-transform technique is noise; because the derivative of the phase of the analytical signal is involved in the computation of the IF, the technique is very sensitive to noise in the original signal. In addition, large fluctuations obscure the estimate of IF derived from the Hilbert transform when the amplitude of the envelope of the signal decreases (deBoer and Nuttall, 1997). An alternative technique for estimating instantaneous frequency is to compute the intervals between zero crossings of the signal (Møller, 1977). This technique typically results in a less noisy estimate of IF, but provides fewer measures of IF as a function of time. The two methods provide results that are generally consistent with each other, as shown below (Fig. 2). The zero crossings provided a simple description of the major trends in the IF as a function of time that was less prone to fluctuations caused by either noise in the signal or changes in the envelope amplitude.

The IF estimated using the intervals between zero crossings was susceptible to dc bias introduced in the waveform before determination of zero crossings, or indeed, to alterations of the dc content of the signal that were introduced by smoothing techniques. In order to minimize the contribution of dc fluctuations (which are unavoidable over short-time intervals of amplitude-modulated signals), the following procedure was used: The time range to be analyzed was determined by computing the envelope of the revcor function (Appendix B) and finding the time range for which the amplitude of the envelope was within 12 dB of the peak of the envelope (deBoer and Nuttall, 1997). The mean (dc) value was then computed over an integer number of cycles of the best frequency centered within this time range. This mean value was then subtracted from the waveform before zero crossings were detected. Zero crossings were determined by identifying the two time samples in the 50-μs resolution waveform that spanned the crossing, and then applying linear interpolation between the two time samples.

A related technique for computing IF is to measure the distances between maxima and/or minima in the impulse re-

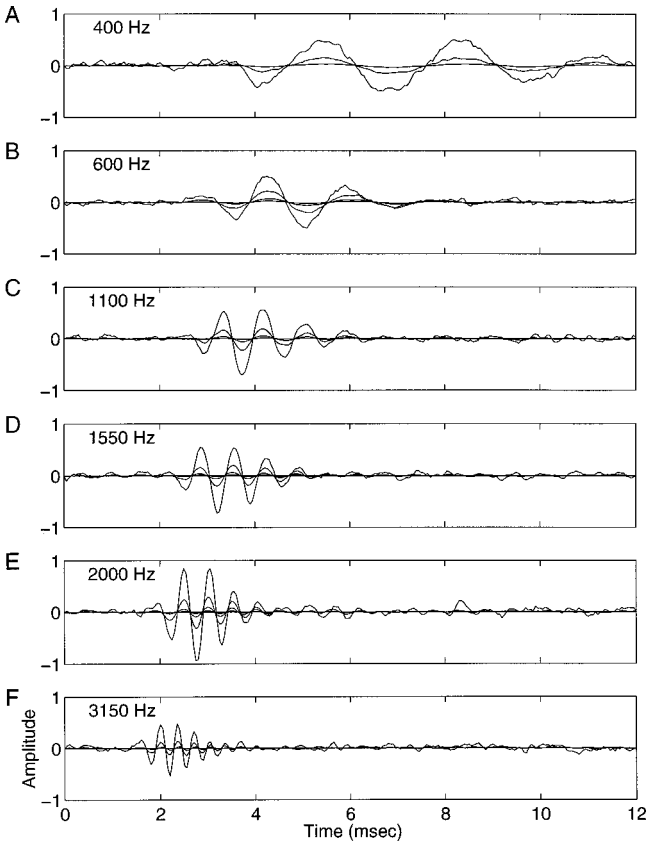


FIG. 1. Revcor functions for six fibers. Revcors for several SPLs are overlaid in each panel, with the higher amplitude waveforms corresponding to the higher SPLs. The BF computed from the revcor function at the lowest SPL is indicated in each panel. Amplitudes are in arbitrary units; all waveforms were normalized by the same value here and in all subsequent figures. These waveforms illustrate the revcor functions after a simple 3-pt smooth was applied; the octave-band filter used in subsequent analysis was not applied to the waveforms shown here. (A) Unit 86204u3; 60, 70, 80 dB SPL. (B) Unit 86100u8; 50, 60, 70, 80 dB SPL. (C) Unit 86100u24; 50, 60, 70, 80 dB SPL. (D) Unit 86065u5; 20, 30, 40, 50, 60, 70, 80 dB SPL. (E) Unit 86100u25; 30, 40, 50, 60, 70, 80 dB SPL. (F) Unit 86204u28; 50, 60, 70, 80 dB SPL. The same waveforms were used for the analyses of instantaneous frequency illustrated in Fig. 3, with different symbols for each SPL.

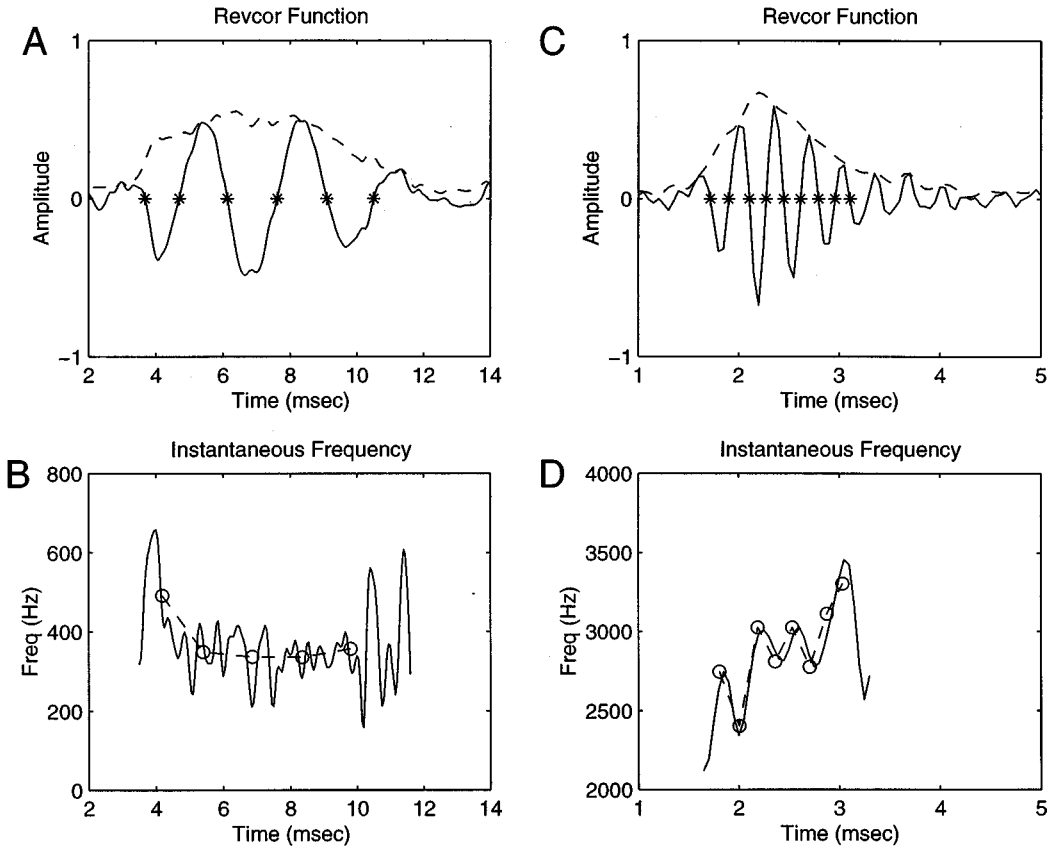


FIG. 2. (A) Revcor function for an AN fiber with BF=400 Hz [same as in Fig. 1(A)]. The dashed line is the envelope computed with the Hilbert transform (Appendix B). A 3-point smooth was applied to the envelopes to reduce the influence of high-frequency noise on the determination of the temporal window to be analyzed (see methods). Asterisks indicate zero crossings. (B) Same as (A), for an AN fiber with BF=3150 Hz [same fiber as in Fig. 1(F)]. (C), (D) IF for each fiber computed using the Hilbert transform (solid lines) and using zero crossings (circles connected by dashed line). Circles are positioned along the time axis midway through each zero-crossing interval.

sponse. This approach is not susceptible to changes in dc bias of the waveform; however, it is susceptible to high-frequency noise in the vicinity of the maxima and minima. Results using the latter approach were consistent with the dc-compensated zero-crossing approach for the majority of the fibers; however, the IFs based on maxima and minima were particularly noisy for the very low-BF fibers; thus, this approach was not used for the results presented here.

II. RESULTS

Revcor functions for six fibers with a range of BFs are illustrated in Fig. 1. Each panel shows revcor functions for a single fiber computed from noise responses at several SPLs; waveforms with higher amplitude peaks correspond to responses to higher amplitude stimuli. The amplitude of the revcor functions[2] (expressed in arbitrary units) in this figure, and all other figures, was normalized by the same value to allow comparison across plots. The fibers illustrated here were arbitrarily chosen from fibers that were studied at a number of SPLs; the properties of these responses were typical for the revcor functions in this study.

Changes in the distances between zero crossings can be observed directly in some of these plots. For the lowest BF fiber [Fig. 1(A)], the first two zero crossings were clearly more closely spaced than subsequent ones. Changes in the zero crossings as a function of time were much smaller for the intermediate BF fibers. For the highest BF fiber [Fig. 1(F)], the zero crossings became more closely spaced as a function of time. Close scrutiny of previously reported revcor functions (e.g., Evans, 1985) and click responses (Kiang *et al.*, 1965) for fibers with different BFs also reveal changes over time in the intervals between zero crossings or intervals between histogram peaks that are similar to those shown in Fig. 1.

A striking feature of the revcor functions shown in Fig. 1 is that the zero crossings were nearly unchanged over a wide range of SPLs. This fact is consistent with earlier reports of AN revcors at several SPLs (Carney and Yin, 1988) and with click responses measured either in the AN (Kiang *et al.*, 1965) or basilar membrane (Recio *et al.*, 1998). The consistency of the zero crossings over a wide range of SPLs indicates that the instantaneous frequency, whether it changed or remained stationary over time, would be similar across SPLs.

The analysis of instantaneous frequency (IF) is illustrated in Fig. 2. Panels A and C show the revcor functions for two fibers at 80 dB SPL [same fibers as in Fig. 1(A),(F)]. The envelope of the revcor function, as determined by the Hilbert transform (Appendix B) is shown as the dashed line, and zero crossings are indicated by asterisks [Fig. 2(A),(C)]. The IF computed using the Hilbert transform is indicated as

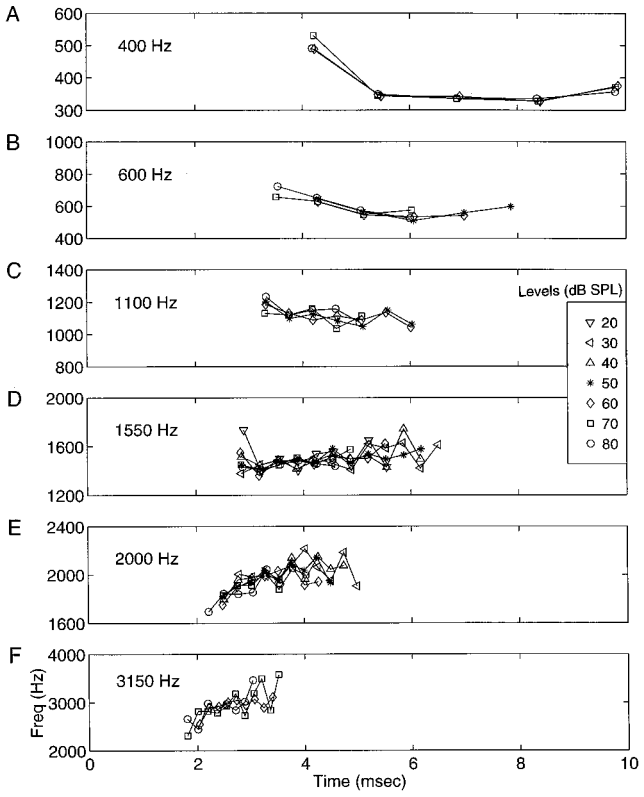


FIG. 3. IF trajectories as a function of time for the six fibers in Fig. 1, at several SPLs. Symbols for SPLs in the legend apply to all plots. IF was computed from zero crossings, as illustrated in Fig. 2. The BF of each fiber, computed from the revcor function at the lowest SPL, is indicated on each plot.

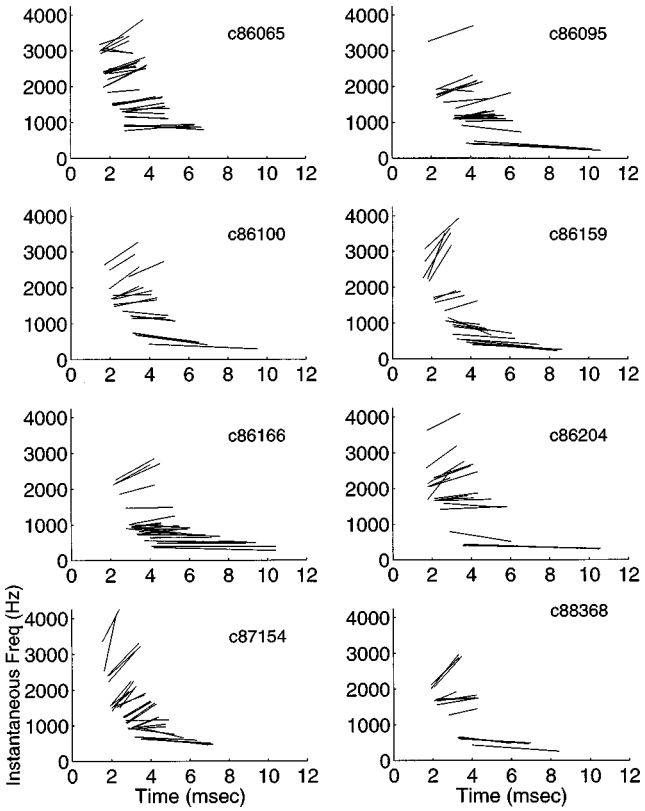


FIG. 5. 1st-order regression lines through IF trajectories for groups of fibers from eight cats. All fibers studied in each cat are included in each plot. IF trajectories were based on zero crossings. All responses were at 80 dB SPL.

the solid line [Fig. 2(B),(D)]. These functions tended to be noisy, as described in the methods, and they were particularly sensitive to noise when the amplitude of the envelope dropped to low values (at the beginning and end of each waveform). The IF computed on the basis of zero crossings is illustrated by the circles and dashed lines. The IF computed from zero crossings was quite consistent with the IF from the Hilbert transform [Fig. 2(D)], including some of the fluctuations over time that were characteristic of especially higher-BF results from the basilar membrane (see deBoer and Nuttall, 1997).

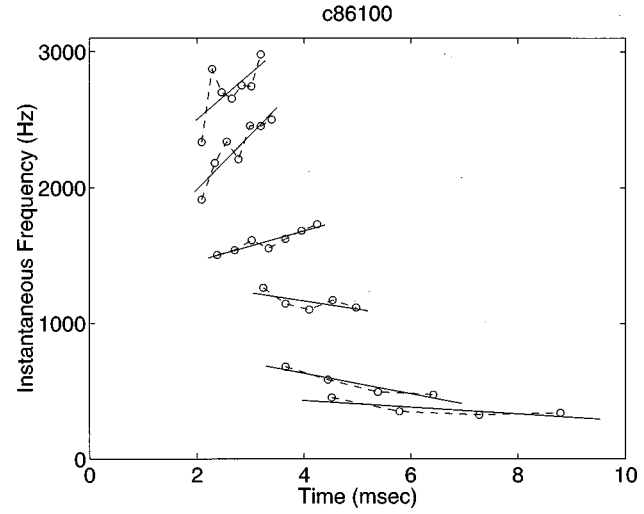


FIG. 4. IF trajectories at 80 dB SPL for six fibers from one cat. IFs were determined using zero crossings (circles connected by dashed lines). Solid lines are 1st-order regressions fit through each trajectory. Units 2, 7, 18, 20, 22, and 26 from cat 86100.

Figure 3 illustrates the IFs for the six fibers in Fig. 1, each computed at several SPLs using zero crossings. The lowest BF fiber [Fig. 3(A)] had an IF that decreased as a function of time, whereas the highest BF fiber [Fig. 3(F)] had an increasing frequency glide. Intermediate BFs had smaller changes in IF as a function of time, progressing gradually from negative slopes at very low BFs to positive slopes at higher BFs. Consistent with the nearly invariant zero crossings at different SPLs (Fig. 1), the IF trajectories for each fiber (Fig. 3) were very similar over the range of SPLs studied here.

Figure 4 illustrates IF trajectories based on zero crossings for several fibers from one cat at one SPL (80 dB SPL). Each trajectory was fit by a first-order linear regression. These regressions and their slopes will be reported in the following figures. For nearly all fibers studied, a 1st-order regression provided a reasonable representation of the IF glide. Some fibers had glides that were not linear functions of time; for example, several (but not all) of the very low-BF fibers had a bend in the IF trajectory [see the two lowest BF fibers in Fig. 4 and the fiber in Fig. 2(B) and Fig. 3(A)]. However, higher-order regressions did not improve the characterization of the major trends in the glides across the population of fibers studied. Furthermore, there was no improve-

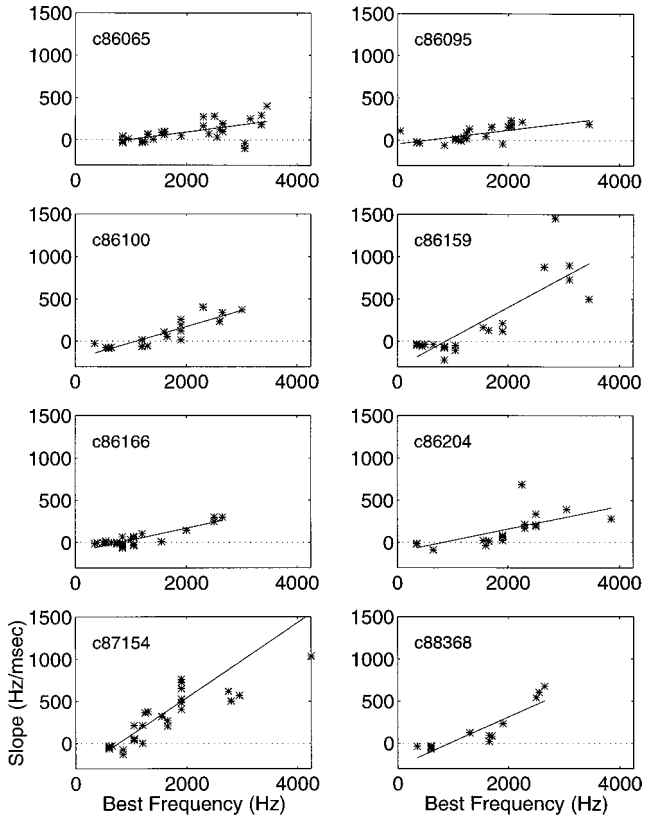


FIG. 6. Slopes of the trajectories in Fig. 5 as a function of best frequency (BF). BF was estimated by the peak of the fast Fourier transform of the revcor function. Each asterisk indicates the slope of a single regression line in Fig. 5. The solid lines are first-order regressions for each set of data points.

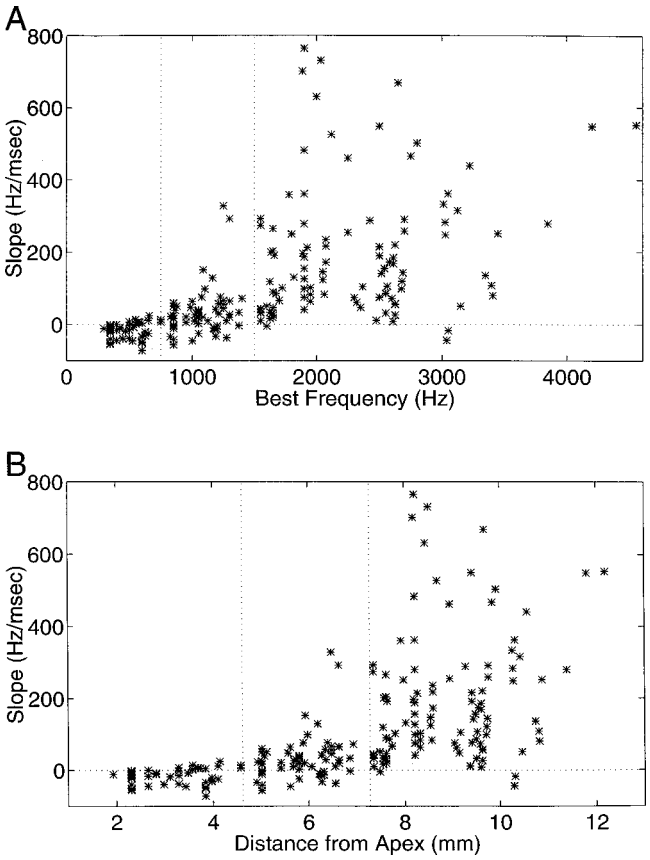


FIG. 7. Slope of the 1st-order regressions of IF trajectories based on responses at all SPLs studied for 214 fibers in 13 cats. The BF for each fiber was determined by averaging the BFs across the SPLs studied. (A) Results plotted on a linear frequency axis. (B) Results plotted on a place axis, using Liberman's (1982) expression for place as a function of CF. Vertical dotted lines are positioned by eye at 750 and 1500 Hz to illustrate trends over different frequency regions.

ment in characterizing these glides by representing them as functions of $\log(t)$, which was the form of the gammachirp suggested by Irino and Patterson (1997). Functions of this type could not provide a simple representation for IF glides across the range of BFs, because the glides vary significantly across BF in direction and overall latency.

Figure 5 illustrates the same characterization as in Fig. 4 for the eight cats in which several fibers across a range of BFs were studied. All responses shown in this figure are for 80-dB SPL noise stimuli. Only the 1st-order regression lines are shown to reduce the complexity of the figure. The major trend in the data is apparent in all cats: positively sloping IF glides at high BFs, flat trajectories at BFs near 1000 Hz, and negatively sloping IF glides at very low BFs.

Figure 6 shows slopes of the trajectories in Fig. 5 as a function of best frequency (BF). Each asterisk indicates the slope of a single regression line in Fig. 5. The slope of the IF trajectories (in Hz/ms) increased as a function of BF in all cats. Slopes were typically positive and relatively large for high BFs, small for BFs near 1000 Hz , and negative for very low BFs.

Trends in the slope of the IF as a function of BF are also apparent in Fig. 7, which illustrates the slopes for 214 fibers from all 13 cats. For the results shown in this figure, the IF was estimated at all SPLs for each fiber. Responses to SPLs that yielded a clean IF (e.g., more than three zero crossings and at SPLs high enough to yield detectable revcor functions) were included in this population plot. IFs for all SPLs were used to determine the 1st-order regression lines; the average BF across the SPLs studied for each fiber was used to determine the position along the BF axis.

Figure 7(A) illustrates the slope of the 1st-order regression for each fiber, plotted in terms of linear frequency. Figure 7(B) plots the slope of the IF regression as a function of distance along the basilar membrane, using Liberman's (1982) function for the relationship between CF and place along the cochlea. The latter plot illustrates the spatial range along the basilar membrane in the cat for which the IF trajectories are negative or relatively flat. Based on previous reports of increasing glides for high-BF fibers and basilar-membrane (BM) responses, it would be assumed that the IF trajectories remain positive for all basilar-membrane positions further from the apex (i.e., BFs higher than 4 kHz).

The vertical dotted lines in each plot (Fig. 7) were positioned at 750 and 1500 Hz to demarcate the three interesting regions of BFs and IF slopes. For fibers with BFs below 750 Hz, 78% of the slopes of the IF trajectories were negative (38 out of 49 fibers), and the positive slopes were quite small (all below 25 Hz/ms). For fibers with BFs between 750 and 1500 Hz, 70% of the IF trajectories had positive slopes (43 out of 61 fibers). Again, these slopes were relatively small: 93% had slopes less than 100 Hz/ms. Above 1500 Hz, the IF trajectories of nearly all (97%, 101 out of 104) fibers

had positive slopes. The size of the slopes varied a great deal for the high-BF fibers; some of this variability was due to combining responses across cats. The results in Fig. 5 for individual cats showed less variability within each animal; Fig. 5 also showed large differences in the absolute values of the high-BF slopes across animals.

III. DISCUSSION

These results show that the frequency modulations, or glides, in the impulse responses of low-frequency AN fibers ($<$1500 Hz) are inherently different from those at high BFs. Whereas the high-BF AN and basilar-membrane responses reported previously all had glides that increased in frequency as a function of time, the direction of the glides is the opposite at very low frequencies. Furthermore, in the mid-frequency region (750–1500 Hz), the glides are relatively flat.

The available database did not include responses to very high SPL stimuli, so it cannot be determined whether the glides would remain unchanged at very high SPLs. Recently, Lin and Guinan (1998) reported significant changes in the impulse responses of AN fibers in the cat in response to very high-level clicks, consistent with a change in the nature of the mechanical input to the hair cells at very high levels. More work is required to determine whether similar changes would be found in indirect estimates of impulse responses based on noise responses.

As pointed out for measurements of high-BF AN fibers and places (deBoer and Nuttall, 1997; Recio *et al.*, 1997, 1998), the trajectories of IF do not change significantly with SPL over the levels studied here. This result is interesting for many reasons. First, it indicates that the glides are not a property due to nonlinearities of the inner ear, such as the active process, which varies the gain and bandwidth of tuning as a function of SPL over the range of levels studied here. Second, impulse responses that have nearly unchanged zero crossings as a function of level, for which there is now an abundance of evidence in both basilar-membrane and AN data, are not consistent with most existing models for cochlear mechanics (deBoer and Nuttall, 1997). Whereas phenomenological models, based on linear filters with carefully placed poles, can exhibit glides (deBoer and Nuttall, 1996, 1997; Shekhter and Carney, 1997), the challenge is to understand a nonlinear system with gain and bandwidth that change with SPL, but with IF trajectories that remain unchanged. A framework for considering the input–output cross-correlation functions for nonlinear systems, and for relating them to linear systems with equivalent cross-correlation functions, has recently been outlined in deBoer's EQ-NL theory (deBoer, 1997). Thus, according to that theory, the results presented here suggest that there must also be a related class of linear systems (as well as nonlinear) that exhibit the nearly invariant IF over a wide range of SPLs, while varying in other respects (e.g., gain and bandwidth).

The constraints that the increasing IF glides for mid- to high BFs have placed on cochlear models have been recognized previously (Møller and Nilsson, 1979; deBoer and Nuttall, 1997). The additional evidence that the slope and direction of the glides changes as a function of BF places new constraints on these models. An increasing frequency glide is not an inherent property of the impulse responses of AN fibers across all BFs. Either the properties of the cochlear mechanics that contribute to the glide must vary along the length of the cochlea, or interactions at the low-frequency end of the cochlea, within approximately 6 or 7 mm of the helicotrema [Fig. 7(B)], must result in significant changes in the nature of the IF glides.

Increasing frequency glides have also been hypothesized to be a property of impulse responses for tuning in the human auditory periphery based on psychophysical tuning curves (Irino and Patterson, 1997). The asymmetry of psychophysical tuning curves was modeled by a filter that has a frequency-modulated impulse response, the gammachirp. However, proposed changes in the slope of the gammachirp's IF glide as a function of SPL, which would produce the desired changes in the asymmetry of the filters as a function of SPL, would be inconsistent with this and other reports that the IF glides of the auditory periphery are nearly unchanged as a function of SPL.

Future work will explore the implications for human perception of the IF glides, including their change as a function of BF. The influence of the IF trajectory of the impulse response on the responses to simple and complex sounds will be complicated. The trajectories in IF seen in the impulse response would be hypothesized to influence both the rate and temporal response properties of individual AN fibers, as well as the spatiotemporal response patterns across the population of AN fibers.

ACKNOWLEDGMENTS

This study was based on data collected during research supported by grant NINCDS-NS12732 while working at the University of Wisconsin with Professor Tom Yin. We acknowledge the generous assistance of Dr. Yin and Ravi Kochhar in obtaining the data files and copies of the data notebooks, which had been carefully maintained over the years. At Boston University, this work has been supported by NSF-IBN9601215. We acknowledge the many helpful comments and discussions of this work with our colleagues Egbert deBoer, David Cameron, Michael Heinz, Tai Lin, Christine Mason, Susan Moscynski, and Ling Zheng.

APPENDIX A: INSTRUCTIONS FOR ACCESSING REVCOR FUNCTIONS VIA FTP

The unsmoothed revcor functions for all the fibers included in this study are available via ftp at the address: *engc.bu.edu* Login as *anonymous*, with your user name as the password. Type *cd\pub\andata* to move to the proper directory. Further instructions about the file formats and file names are in the *readme.txt* file in this directory.

APPENDIX B: TIME-DOMAIN ANALYSIS OF REVCOR FUNCTIONS USING THE HILBERT TRANSFORM

A revcor function $h(t)$ can be decomposed into its envelope and instantaneous frequency using the Hilbert transform (e.g., Haykin, 1994). First, an analytic signal is formed

$$a(t) = h(t) + j\hat{h}(t),$$

where $\hat{h}(t)$ is the Hilbert transform of $h(t)$, and $j = \sqrt{-1}$. Then, the envelope of the revcor function env(t) is equal to the magnitude of the complex function $a(t)$

$$\text{env}(t) = |a(t)|,$$

while the instantaneous frequency (IF, in Hz) is proportional to the derivative of the phase angle of $a(t)$

$$\text{IF}(t) = [d/dt\{\angle a(t)\}]/(2\pi).$$

[1]In this manuscript, the term best frequency (BF) will be used to describe the center frequency of tuning of AN fibers and places along the basilar membrane. BF is most often used to describe tuning based on suprathreshold stimuli, whereas characteristic frequency (CF) is generally used to describe tuning for threshold level stimuli. At low stimulus levels, estimates of BF based on revcor functions are approximately equal to CF (deBoer, 1973). Here, for simplicity, the term BF will be used throughout, because the measures presented here are all based on suprathreshold stimuli.

[2]The amplitude of a revcor function is influenced by the number of action potentials, the degree of phase-locking to energy near BF within the noise stimulus, and the amplitude and units of the stimulus waveform used in the analysis (e.g., the waveform may be expressed in volts or Pa).

Boer, E. de (**1967**). ''Correlation studies applied to the frequency resolution of the cochlea,'' J. Aud. Res. **7**, 209–217.

Boer, E. de (**1973**). ''On the principle of specific coding,'' J. Dyn. Syst., Meas., Control **95G**, 265–273.

Boer, E. de (**1997**). ''Connecting frequency selectivity and nonlinearity for models of the cochlea,'' Aud. Neurosci. **3**, 377–388.

Boer, E. de, and Jongh, H. R. de (**1978**). ''On cochlear encoding: Potentialities and limitations of the reverse correlation technique,'' J. Acoust. Soc. Am. **63**, 115–135.

Boer, E. de, and Kuyper, P. (**1968**). ''Triggered correlation,'' IEEE Trans. Biomed. Eng. **15**, 169–179.

Boer, E. de, and Nuttall, A. L. (**1996**). ''Cochlear travel time and minimum phase,'' Assoc. Res. Otolaryngol. **19**, 57.

Boer, E. de, 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.

Carney, L. H., and Yin, T. C. T. (**1988**). ''Temporal coding of resonances by low-frequency auditory nerve fibers: Single fiber responses and a population model,'' J. Neurophysiol. **60**, 1653–1677.

Carney, L. H. (**1990**). ''Sensitivities of cells in the anteroventral cochlear nucleus of cat to spatio-temporal discharge patterns across primary afferents,'' J. Neurophysiol. **64**, 437–456.

Carney, L. H. (**1993**). ''A model for the responses of low-frequency auditory nerve fibers in cat,'' J. Acoust. Soc. Am. **93**, 401–417.

Eggermont, J. J. (**1993**). ''Wiener and Volterra analyses applied to the auditory system,'' Hearing Res. **66**, 177–201.

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.

Evans, E. F. (**1981**). ''The dynamic range problem: Place and time coding at the level of cochlear nerve and nucleus,'' in *Neuronal Mechanisms of Hearing*, edited by J. Syka and L. Aitkin (Plenum, New York), pp. 69–86.

Evans, E. F. (**1985**). ''Aspects of the neural coding of time in the mammalian peripheral auditory system relevant to temporal resolution,'' in: *Time Resolution in Auditory Systems*, 11th Danavox Symposium, edited by A. Michelsen (Springer, New York), pp. 74–95.

Haykin, S. S. (**1994**). *Communication Systems* (Wiley, New York).

Irino, T., and Patterson, R. D. (**1997**). ''A time-domain, level-dependent auditory filter: The gammachirp,'' J. Acoust. Soc. Am. **101**, 412–419.

Kiang, N. Y.-S., Watanabe, T., Thomas, E. C., and Clark, L. F. (**1965**). ''Discharge patterns of single fibers in the cat's auditory nerve,'' MIT Research Monograph No. 35 (MIT, Cambridge, MA).

Lee, Y. W., and Schetzen, M. (**1965**). ''Measurement of the Wiener kernels of a non-linear system by cross-correlation,'' Int. J. Control **2**, 237–254.

Liberman, M. C. (**1982**). ''The cochlear frequency map for the cat: labelling auditory nerve fibers of known characteristic frequency,'' J. Acoust. Soc. Am. **72**, 1441–1449.

Lin, T., and Guinan, Jr., J. J. (**1998**). ''Auditory-nerve discharge patterns in response to high-sound-level clicks suggest that two different resonances are involved,'' Assoc. Res. Otolaryngol. **21**, 137.

Marmarelis, P. Z., and Marmarelis, V. Z. (**1978**). *Analysis of Physiological Systems: the White Noise Approach* (Plenum, New York).

Møller, A. R. (**1977**). ''Frequency selectivity of single auditory-nerve fibers in response to broadband noise stimuli,'' J. Acoust. Soc. Am. **62**, 135–142.

Møller, A. R. (**1981**). ''Coding of complex sounds in the auditory nervous system,'' in *Neuronal Mechanisms of Hearing*, edited by J. Syka and L. Aitkin (Plenum, New York), pp. 87–103.

Møller, A. R. (**1983**). *Auditory Physiology* (Academic, New York), pp. 220ff and Fig. 3.13.

Møller, A. R., and Nilsson, H. G. (**1979**). ''Inner ear impulse response and basilar membrane modelling,'' Acustica **41**, 258–262.

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 vibrations of the basilar membrane in squirrel monkeys using the Mössbauer technique,'' J. Acoust. Soc. Am. **49**, 1218–1231.

Robles, L., Rhodes, W. S., and Geisler, C. D. (**1976**). ''Transient response of the basilar membrane measured in squirrel monkeys 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,'' J. Acoust. Soc. Am. **80**, 1364–1374.

Rosen, S., and Baker, R. J. (**1994**). ''Characterizing auditory filter nonlinearity,'' Hearing Res. **73**, 231–243.

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.

Shekhter, I. (**1997**). ''A phenomenological model for nonlinear response properties of auditory-nerve fibers,'' Masters thesis, Boston University.

Shekhter, I., and Carney, L. H. (**1997**). ''A nonlinear auditory nerve model for CF-dependent shifts in tuning with sound level,'' Assoc. Res. Otolaryngol. **20**, 617.

Maturation of medial efferent system function in humans

Carolina Abdala,[a] Ellen Ma, and Yvonne S. Sininger
House Ear Institute, Children's Auditory Research and Evaluation Center, 2100 West Third Street, Los Angeles, California 90057

(Received 13 October 1998; accepted for publication 21 January 1999)

Otoacoustic emissions are typically reduced in amplitude when broadband noise is presented to the contralateral ear. This contralateral suppression is attributed to activation of the medial olivocochlear system, which has an inhibitory effect on outer hair-cell activity. By studying the effects of contralateral noise on cochlear output at different stages of auditory maturation in human neonates, it is possible to describe the timecourse for development of medial efferent system function in humans. The present study recorded $2f1-f2$ distortion product otoacoustic emissions (DPOAE) in human adults, term and premature neonates at three $f2$ frequencies: 1500, 3000, and 6000 Hz, using fixed primary tone frequency ratio ($f2/f1=1.2$) and level separation (10 dB, $L1>L2$). Average DPOAE growth functions were recorded with and without contralateral broadband noise. Results indicate that contralateral suppression of DPOAEs is absent at 6000 Hz, but present at 1500 and 3000 Hz for all ages. However, DPOAE amplitude from premature neonates was not altered by noise in an adult-like manner; in this age group, DPOAE amplitude was equally likely to be suppressed or *enhanced* by noise presented contralaterally. Contralateral enhancement may reflect a temporary stage of immaturity in outer hair cell-medial efferent fiber synapses just prior to term birth. © *1999 Acoustical Society of America.* [S0001-4966(99)04504-X]

PACS numbers: 43.64.Jb, 43.64.Kc, 43.64.Ri [BLM]

INTRODUCTION

The medial olivocochlear (MOC) pathway courses from mid-brainstem to innervate primarily outer-hair cells (OHC) of the cochlea. Most efferent innervation to OHCs comes from the superior olivary complex opposite the targeted ear to form the crossed MOC bundle. The uncrossed pathway is smaller and courses from the ipsilateral complex to the outer-hair cells. The exact function of these efferent fibers is not known. In general, the MOC system has an inhibitory effect on the auditory periphery, although some observations of enhancement have also been reported (Jenkins *et al.*, 1993; Kawase and Liberman, 1993; Liberman *et al.*, 1996; Siegel and Kim, 1982). Because of its predominantly inhibitory nature, it has been hypothesized that the efferent system serves a protective role in the auditory system (Rajan, 1990). It is also hypothesized that the activation of the efferent system enhances the detection of sound in noise (Micheyl and Collet, 1996; Winslow and Sachs, 1987) and maintains the cochlea at an optimum mechanical state for efficient function of active processes (Johnstone *et al.*, 1986).

Equivocal results have been reported describing the effects of severing the olivocochlear bundle on auditory coding and processing. Some reports in nonhuman species find degraded frequency resolution when MOC fibers are disrupted (Capps and Ades, 1968; Carlier and Pujol, 1982). Others have found an unchanged auditory process after eliminating MOC influence (Igarishi *et al.*, 1979; Littman *et al.*, 1992; Rajan, 1990). In humans, patients receiving vestibular neurotomy have shown no detectable impairment in hearing processes (Scharf *et al.*, 1994). Recent work in the cat suggests that degradation of auditory function after eliminating efferent influence is only observed if disruption occurs during critical developmental periods (Walsh *et al.*, 1998).

Although the biological significance of medial efferent fibers remains ambiguous, it is clear that activation of medial olivocochlear fibers alters cochlear output. For example, electrical stimulation of the MOC bundle at the floor of the fourth ventricle inhibits cochlear and neural potentials (Galambos, 1956; Guinan and Gifford, 1988; Mountain, 1980; Siegel and Kim, 1982). MOC fibers can also be activated by acoustic stimulation. Broadband noise presented to the ear opposite the test ear reduces amplitude of several auditory responses, such as the compound action potential (Folsom and Owsley, 1987; Liberman, 1989), spontaneous rate of auditory-nerve fibers (Buno, 1978), and otoacoustic emissions (Berlin *et al.*, 1994; Collet *et al.*, 1990; Puel and Rebillard, 1990). Upon section of the MOC bundle, contralaterally presented broadband noise becomes ineffective in altering cochlear output (Liberman, 1989; Puel and Rebillard, 1990; Puria *et al.*, 1996). In the last decade, contralateral suppression of otoacoustic emissions has been conducted in laboratory animals, human adults (Mott *et al.*, 1989; Puel and Rebillard, 1990; Liberman *et al.*, 1996; Berlin *et al.*, 1994; Collet *et al.*, 1990), and most recently in human neonates as a measure of efferent system maturation (Abdala *et al.*, 1998; Morlet *et al.*, 1993; Ryan and Piron, 1994).

Because otoacoustic emissions (OAE) are a normal by-product of cochlear amplifier activity and reflect OHC integrity, they provide an appropriate index of changes in cochlear function as MOC fibers are activated. In human adults, transient-evoked OAEs (TEOAE) show an average reduction of 3.7 dB with presentation of moderate levels of

[a]Author to whom correspondence should be addressed. Electronic mail: Cabdala@hei.org

broadband noise in the contralateral ear (Veuillet *et al.*, 1991). Contralateral suppression of distortion-product OAEs has been minimally studied in humans. Most reports show an average of 0.5 to 2 dB reduction in DPOAE amplitude with noise in the contralateral ear, depending on level and frequency of primary tones (Abdala *et al.*, 1998; Moulin *et al.*, 1993; Williams and Brown, 1997).

If the effects of broadband noise on OAEs are measured at different stages of maturation, it can provide a window into development of medial efferent system function in humans. Only two published studies have investigated contralateral suppression in neonates. Both recorded transient-evoked OAEs (TEOAEs). Morlet and colleagues (1993) found that broadband noise presented contralaterally had no effect on TEOAE amplitude for a group of premature neonates ranging in conceptional age from 33 to 39 weeks at test (Morlet *et al.*, 1993). Other reports have observed contralateral suppression of TEOAEs in term-born neonates and even in some premature subjects (Goforth *et al.*, 1997; Ryan and Piron, 1994).

Contralateral suppression of *distortion-product* OAEs has not been investigated in human neonates, yet it provides a noninvasive and frequency-specific probe of human medial efferent system function. The objective of this study was to study DPOAE contralateral suppression in adults, term and premature neonates to investigate the frequency- and level-dependent maturation of the medial efferent system function in the human auditory system.

I. METHODS

A. Subjects

Twenty-four ears (12 left and 12 right) from 15 normal-hearing adults were tested with the experimental protocol. The adults had a mean age of 28 years and had audiometric thresholds <15 dB between 250 and 8000 Hz. Seven of the adult subjects were male and eight female. None of the adult subjects had a history of noise exposure or otologic disease.

Fifty-two term-born and 35 premature neonates were included as subjects in this study. The term-born neonates were born between 37 and 41 weeks gestation, averaged 3393 grams at birth, and had average 5- and 10-min APGAR scores of 8.3 and 8.8, respectively. Just over 50% of the term neonates were tested within 72 h of birth. The other half were tested no more than 7 days after birth. Availability for test depended on many factors, including the health status of the mother and procurement of an informed consent.

The 35 premature subjects were born between 27 and 36 weeks of gestation. 80% of these subjects were tested prior to 37 weeks postconceptional age (CA) (mean age at test $=35$ weeks CA). Thus, the preterm subjects were not only prematurely born, but most remained premature at the time of our evaluation. See Fig. 1(a) and (b) for a breakdown of ages at birth and test for premature subjects. Average birthweight for premature subjects was 2095 grams, and average 5- and 10-min APGARs were 7.1 and 8.7, respectively.

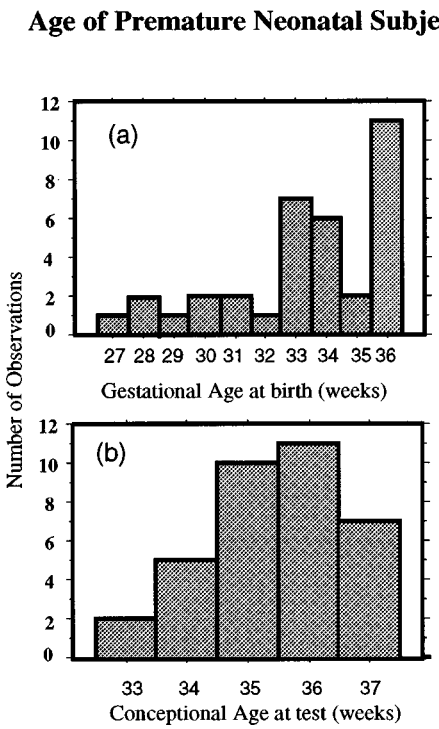


FIG. 1. (a) Gestational age of premature neonatal subjects at birth; (b) Conceptional age of subjects at test. Conceptional age was calculated by adding gestational age at birth to weeks between birth and test.

B. Instrumentation and signal analysis

An Ariel DSP16+ signal processing and acquisition board housed within a Compac Prolinea 590 personal computer with Pentium processor was used to generate stimuli and acquire data. The Ariel board was connected to an Etymotic Research ER-10C probe-mic system. The ER-10C probe contains two output transducers and a low-noise microphone. The two primary tones were generated by the DSP processor.

Energy at the probe microphone was high-pass filtered (12 dB/oct: 710 Hz high-pass cutoff) and sampled at a rate of 50 kHz with a sweep length of 4096 samples, giving a frequency resolution of 12.2 Hz. Twenty-five sweeps of the microphone signal were added and comprised one block for $f2=3000$ and 6000 Hz. Due to elevated noise in the low frequencies, 50 sweeps were added to make up one block at 1500 Hz. The sum of each block was stored as 32-bit integer values. Two acceptable blocks of data were summed (see the data acceptance criteria below), averaged, and the power spectrum was obtained by applying a 4096-point FFT. The level of the DPOAE was computed using the summed 32-bit data directly and 8-byte IEEE-format floating point computations in the discrete Fourier transform.

Intermodulation distortion produced by the recording system at $2f1-f2$ was measured with the probe in a Zwislocki coupler for all test conditions. The mean level of system distortion was -21 dB SPL. In no case did the level exceed -17 dB SPL. The recording system noise floor was determined using a similar method with no tones present. The level of system noise floor ranged between -22 and -27 dB SPL, depending on frequency.

For adult subjects, the broadband noise (BBN) presented

to the contralateral ear was generated by a Grason–Stadler 16 audiometer in the white noise modality. For neonatal subjects, white noise was generated by the STIM program of a Neuroscan evoked potentials system. Noise was delivered through an Etymotic Research, ER-2 transducer for both adults and neonates.

C. Data acceptance criteria

DPOAE measurements were accepted and recorded as valid data if they met the following criteria: (1) Noise measurements for three frequency bins on either side of the $2f1-f2$ frequency had to be <0 dB SPL to ensure appropriate subject state. Preliminary work has shown that subjects with noise of 0 dB SPL or lower are more likely to generate DPOAE data with low amplitude variability and, (2) The measured DPOAE level was at least 5 dB above the average noise measured in the same six bins around the DPOAE frequency.

The program attempted up to six blocks of either 25 or 50 sweeps to obtain two acceptable blocks. If two blocks were not collected after six attempts, no data were collected and the next condition was initiated. In addition, sweeps were accepted into a block of data only when the estimated rms level in that sweep did not exceed a user-controlled artifact rejection threshold. This level was set for each subject based on observations of baseline activity level determined early in the test session, and modified if necessary during the experiment.

D. Procedure

Adult subjects were tested in a sound-treated booth at the House Ear Institute while reading, sleeping, or sitting quietly. Infant subjects were evaluated at the Infant Auditory Research Laboratory, Los Angeles County+University of Southern California Medical Center, Women and Children's Hospital, after informed consent was obtained from parents. Neonates were tested in an acoustically treated infant isolette (Eckels ABC-100 acoustic isolette) that provided 25 dB of attenuation from 500 to 2000 Hz and up to 45 dB of attenuation between 2000 and 8000 Hz. The neonates were fed if necessary prior to the test, swaddled, and placed in the isolette to sleep.

An initial *in situ* calibration procedure was conducted on both output transducers before each subject was tested. Tones of fixed voltage were presented to the transducers at 250-Hz intervals from 500 to 10 000 Hz and the resulting SPL of these tones was recorded in the ear canal. Based on this information, an equalization of output levels was performed for each subject to achieve target stimulus and suppressor levels across all test frequencies. Three $f2$ frequencies (1500, 3000, and 6000 Hz) were presented with 10-dB separation between $L1$ and $L2$ $(L1>L2)$ and a constant $f2/f1$ ratio of 1.22. DPOAE amplitude was recorded at 12 stimulus levels ranging from 30 to 85 dB SPL in 5-dB intervals. Three growth functions were collected with BBN (+BBN) and three without BBN (−BBN) in an alternating fashion, for a total of six growth functions per ear. Only the mean of three growth functions for +BBN and three for −BBN was entered into data analysis for each ear. This initial averaging of DPOAE amplitude data from any one ear was conducted to reduce the influence of intrasubject, run-to-run variability.

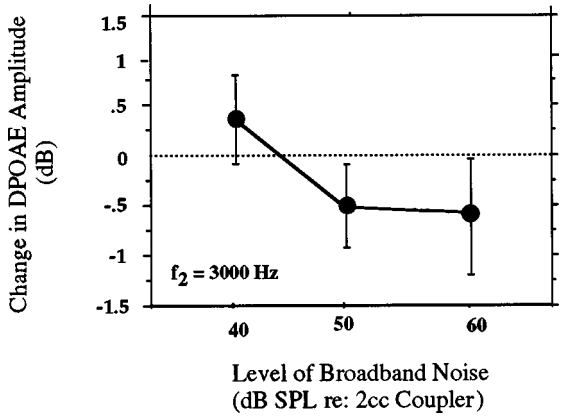


FIG. 2. A pilot group of eight neonates was evaluated to investigate the changes in DPOAE amplitude with contralateral broadband noise (BBN) presented at three different levels. DPOAE amplitude values falling below 0 dB represent a decrease in amplitude with BBN. The noise level is referenced to measurements made in a 2-cc coupler.

E. Broadband noise levels

To determine appropriate noise levels for neonates, DPOAEs were recorded from eight term neonates with 65–55 dB SPL primary tones. BBN was then added to the contralateral ear at 40, 50, and 60 dB SPL (re:2-cc coupler). Figure 2 displays the result of this pilot experiment for $f2$ $=3000$ Hz. BBN presented contralaterally at 40 dB SPL produced little change in the DPOAE amplitude of these neonates, whereas 50 and 60 dB of noise produced comparable decreases in amplitude. Thus, 50 dB SPL was chosen as the appropriate BBN level to evoke DPOAE contralateral suppression in neonates. This level is below levels of noise known to evoke the middle-ear reflex in adults and is in good agreement with previous studies applying this paradigm to humans (Moulin *et al.*, 1993; Silman, 1984; Williams and Brown, 1997). There are no data available describing the acoustic reflex threshold in neonates, and it cannot be unequivocally ruled out that this factor contributed to the changes in DPOAE amplitude observed with presentation of contralateral noise. However, as described in the Results section, the changes noted in neonatal DPOAE amplitude with contralateral noise are not consistent with attenuation provided by middle-ear muscle activity.

Although appropriate BBN level was directly determined for neonates, as described above, the same voltage applied to an infant and adult ear will result in different sound-pressure levels in the ear canal (Sininger *et al.*, 1997; Lasky, 1998). To determine the magnitude of this difference, in-the-ear measurements of BBN were made on a small group of adult and neonatal subjects (Etymotic Research ER-7 probe mic system). With a fixed voltage to the transducer, BBN measured in the ear canal of neonates was on average 14 dB greater than noise levels in adult ear canals. It is assumed that this difference is due to the natural amplification that occurs in a small volume cavity like the neonatal

ear canal. To compensate partially for this factor, BBN was presented to adult subjects at a level 10 dB higher than neonates (i.e., 60 dB SPL *re*: 2-cc coupler).

II. ANALYSIS

A. Noise effects on DPOAE amplitude

Contralateral suppression was initially calculated by subtracting DPOAE amplitude without contralateral noise (−BBN) from DPOAE amplitude in the presence of contralateral broadband noise (+BBN) at each stimulus level. This ''difference score'' provides an index of DPOAE suppression produced by contralateral noise and attributed to activation of the medial olivocochlear system.

One-sample *t*-tests were conducted on contralateral suppression difference scores within age group, at each $f2$ frequency. Because of the elevated noise floor in neonatal subjects, many neonates did not have DPOAE measurements at stimulus levels ($L1$) below 50 dB SPL. Thus, the *t*-tests were conducted at the four primary tone levels with the greatest numbers of observations for all ages ($L1 = 85$, 75, 65, and 55 dB). A Bonferoni adjustment was applied to the *t*-test results.

B. Alternative measures

Two alternative metrics of efferent effect were calculated to more precisely define the effects of contralateral noise and age on DPOAE amplitude:

(1) There were many instances of DPOAE amplitude increase or ''enhancement'' upon presentation of BBN. When calculating a traditional difference score for contralateral suppression, DPOAE amplitude in −BBN is subtracted from amplitude recorded in +BBN. Therefore, if suppression is present, the difference score is a negative number. However, in cases of enhancement, the value is positive. As a result, when contralateral suppression scores are averaged across all subjects, mean values may appear small only due to cancellation that occurs when adding positive and negative difference scores. For this reason, contralateral suppression was recalculated after instances of amplitude enhancement were eliminated. This was done to minimize the impact of DPOAE enhancement when conducting analyses to characterize the suppressive effect.

(2) Because DPOAE contralateral enhancement itself may represent an important efferent effect, it was calculated independently as a function of primary-tone level for the three age groups. This analysis includes only observations of DPOAE amplitude increase upon presentation of contralateral noise.

III. RESULTS

The prevalence of DPOAEs in the neonates selected to participate in this study was 94%. Successful completion of the experimental protocol occurred in 82% of attempted neonatal subjects. In adults, 100% of subjects selected for this study were successfully tested with the protocol. DPOAE growth functions showed generally increasing DPOAE amplitude for low to moderate primary-tone levels and amplitude saturation at moderate to high primary-tone levels. There were a variety of configurations; however, the majority of subjects showed nonmonotonic patterns of DPOAE amplitude growth over the 12 primary-tone levels presented.

A. Effects of contralateral BBN on DPOAE amplitude

Figure 3 displays mean DPOAE amplitude as a function of $L1$ for each age group separately at the three $f2$ frequencies. The parameter is noise condition (+BBN or −BBN). Contralateral suppression was not observed at 6000 Hz for any age group. Thus, subsequent analyses and reporting of results focused on data at 1500 and 3000 Hz only.

Figure 4 displays the mean contralateral suppression difference score as a function of primary tone level at $f2 =$ (a) 1500, (b) 3000, and (c) 6000 Hz. Any mean value falling below the horizontal line at 0 dB indicates contralateral suppression. To better observe age trends, mean contralateral suppression was collapsed across level (Fig. 4-inset). Contralateral suppression at $f2 = 1500$ Hz in adults was significantly greater than 0 dB at all four primary tone levels tested ($p = 0.0007$, 0.0003, 0.0001, and 0.0002). It ranged from 0.5 to 6.6 dB with an average of 1.2 dB. Term infants showed significant contralateral suppression for this same $f2$ frequency at 75 dB SPL ($p = 0.02$). For $f2 = 3000$ Hz, adults had significant DPOAE contralateral suppression at 65, 75, and 85 dB ($p = 0.001$, 0.01, and 0.002) and term infants at 65 and 75 dB ($p = 0.0005$ and 0.0007). Premature neonates did not show significant amounts of DPOAE suppression using the traditional contralateral suppression difference score analysis.

B. Alternative measures of efferent effect

An increase in DPOAE amplitude when noise was presented to the opposite ear was termed ''contralateral enhancement'' and it may reflect an important physiological effect. However, when attempting to specifically characterize suppression (i.e., inhibitory effects of the medial efferent system), episodes of enhancement interfere with and preclude the accurate assessment of the magnitude of suppression when it occurs. Thus, contralateral suppression was calculated a second time after all data points showing contralateral enhancement were eliminated (Fig. 5). As shown in Fig. 5, when DPOAE contralateral enhancement is not averaged into the data, the three age groups overlap significantly at moderate-to-high levels. Contrary to the previous analysis, these data indicate that both term and premature neonates are adult-like in terms of the magnitude of suppression evoked by contralateral noise. Additionally, data in Fig. 5 indicate that suppression increases as primary-tone level decreases for all ages.

Figure 6 shows only contralateral enhancement (all suppression data eliminated) as a function of stimulus level. Enhancement was not observed equally in each age group. Premature neonates as a group showed greater incidence of DPOAE contralateral enhancement: 43% of data from premature neonates showed DPOAE amplitude enhancement upon presentation of BBN. In contrast, only 15% and 24% of adult and term data, respectively, showed enhancement.

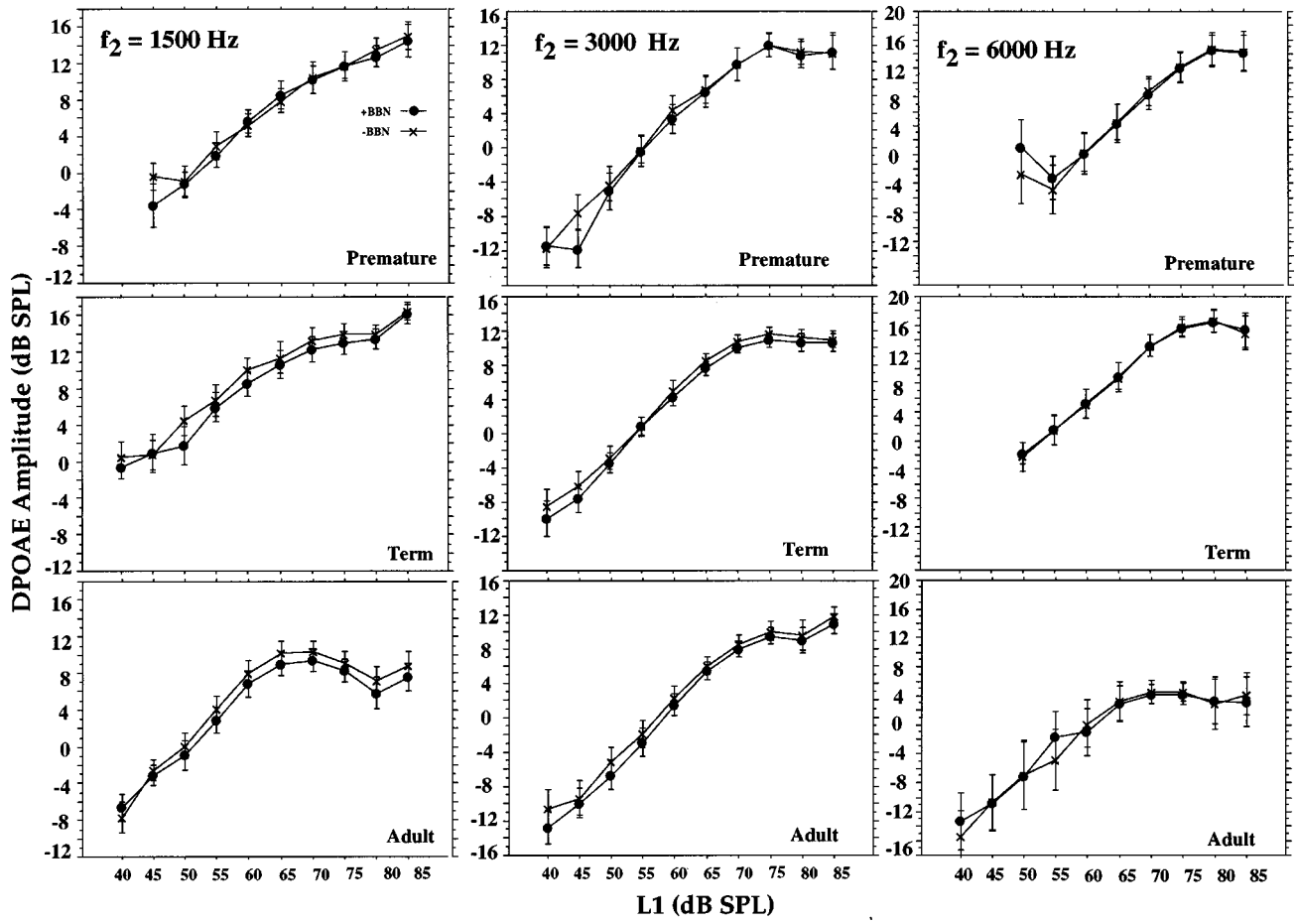


FIG. 3. DPOAE amplitude growth functions recorded in the absence of contralateral noise (−BBN) and with contralateral noise (+BBN) at $f2=1500$, 3000, and 6000 Hz. Data are presented for the three age groups: premature neonates, term neonates, and adult subjects. Each mean is comprised of between five and 32 observations. 70% of the averages were derived from at least ten observations (only three of the averages were derived from five observations).

Only adult and premature neonatal data are included in Fig. 6 in order to display the two extremes of the age variable. The number of ears upon which each mean value is derived is included to the left of the symbol. Considering that 85% of adult data were eliminated for this analysis, the mean values in Fig. 6 are derived from a greatly reduced data set. For this reason, enhancement data are shown at only four stimulus levels for 1500 Hz and five levels for 3000 Hz. These are the levels at which each group had at least three observations to contribute to an average. Although it is difficult to come to a strong conclusion based on such few observations, Fig. 6 suggests that premature neonates have a larger enhancement effect than adults at both $f2$ frequencies.

Because contralateral enhancement is more common in premature neonates than adults (43% vs 15%) and, when present, is a larger effect in premature neonates, the presence of enhancement appears to be related to age. Figure 7(a)–(d) are scatterplots showing the distribution of contralateral suppression and enhancement as a function of conceptional age at test and gestational age at birth for $f2=1500$ and 3000 Hz. Conceptional age at test provides an index of the maturational status of the subject when tested, whereas gestational age at birth provides an index of how premature the subject was when he/she was born. Figure 7 suggests that occurrences of enhancement decrease with increasing conceptional age at 1500 Hz [Fig. 7(a)]. Data in Fig. 7(b) also indicate that there is a relationship between contralateral enhancement at 1500 Hz and gestational age at birth. This suggests that the length of gestation itself may be an important factor in determining medial efferent system maturity and function. There does not appear to be a consistent relationship between conceptional/gestational age and enhancement/suppression at 3000 Hz (data shown) or 6000 Hz.

IV. DISCUSSION

The results of this study show that contralateral noise presented during the recording of ipsilateral $2f1-f2$ DPOAE provides a feasible technique for application to group data of term neonates and very young premature neonates. Variability of DPOAE amplitude was grossly comparable in each age group (see Table I) although neonates generally required longer averaging time to achieve the same signal/noise ratio as adults. This is due to the higher noise floor typically found in neonatal data. Other than the longer test time, success with this technique was considered excellent with neonates and comparable to other DPOAE paradigms used to assess related aspects of cochlear function.

Results of this study show, however, that a simple difference score metric does not appropriately characterize the effect of contralateral noise on DPOAE amplitude in neonates. Because adults and term-born subjects did not show

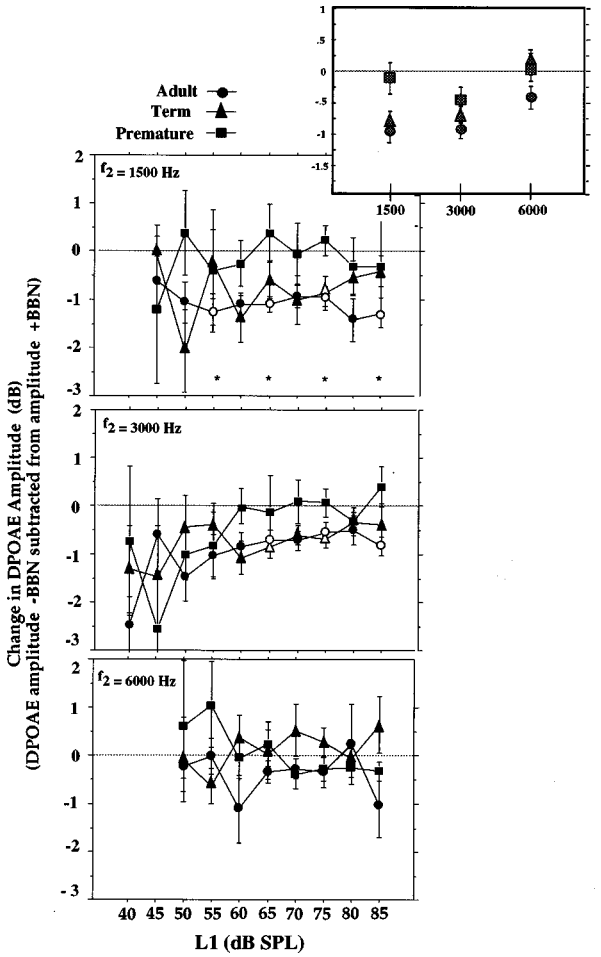


FIG. 4. Mean DPOAE amplitude change is displayed as a function of stimulus level for each of the three age groups at three $f2$ frequencies. Change in amplitude was calculated by subtracting DPOAE amplitude with no contralateral broadband noise (−BBN) from amplitude recorded with contralateral noise (+BBN). Any values falling below 0 dB represent a decrease in amplitude with BBN. Asterisks in the topmost graph ($f2 = 1500$ Hz) denote the stimulus levels at which *t*-tests were conducted for all conditions. Open symbols represent the primary-tone levels at which contralateral suppression was found to be significant.

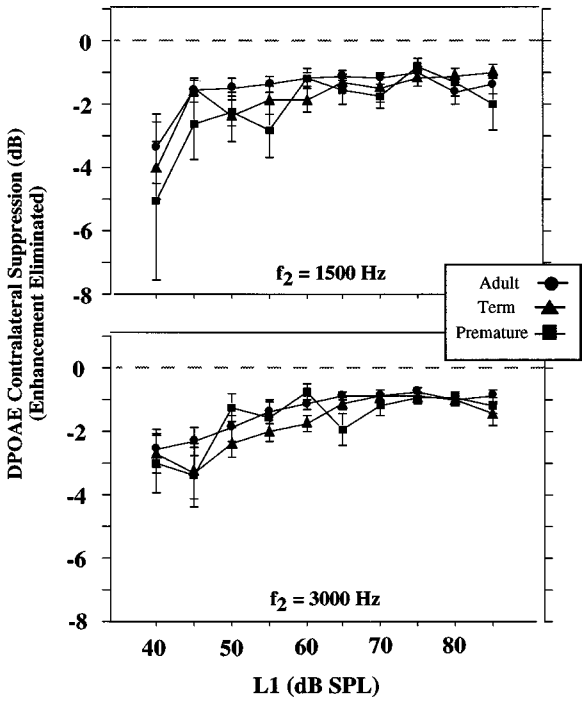


FIG. 5. Mean contralateral suppression is plotted as a function of stimulus level at $f2 = 1500$ and 3000 Hz for three age groups. DPOAE contralateral suppression was calculated by subtracting DPOAE amplitude −BBN from amplitude in +BBN conditions after all data reflecting amplitude enhancement were eliminated. Each mean is comprised of between six and 32 observations; 93% of the averages in this figure are made up of at least ten observations. Vertical bars represent ± one standard error of the mean.

frequent occurrences of DPOAE amplitude enhancement with noise, the traditional calculation of DPOAE contralateral suppression (DPOAE amplitude −BBN subtracted from amplitude +BBN) may be an acceptable tool for analysis in these groups. However, in premature neonates, the effect of noise on the DPOAE is blurred and diluted by averaging enhancement and suppression values together. Thus, if this technique is to be used with neonates for the study of auditory maturation, DPOAE contralateral enhancement and suppression should be calculated independently.

The contralateral BBN/DPOAE paradigm appears to be an appropriate tool for assessing medial efferent system function in *group* data. The application of statistics to evaluate differences in mean data of a group of subjects ensures that the noise effect is tested against normal amplitude variability unrelated to contralateral BBN. Thus, it is possible to establish that contralateral noise effects on DPOAE amplitude exist in group data with a high degree of certainty. However, it is not clear how much of a noise-evoked DPOAE amplitude change must occur in any one individual to consider the effect present or ''normal.'' An effective approach to this problem would be to record three or four abbreviated DPOAE growth functions per individual and then analyze the run-to-run variability. The range of intrasubject variability (measured by standard deviations or confidence intervals) could provide the boundaries for determining the presence of true contralateral noise effects. Similarly, the intrasubject variability derived from a large group of subjects could be used to provide these boundaries. This may be one method for applying this paradigm to individual patients or subjects to assess integrity of medial efferent system function. This issue warrants greater attention if this technique is to be applied in this fashion.

A. The adult medial efferent system

The results of this study indicate that the presentation of contralateral broadband noise reduces output of cochlear OHCs in the healthy cochlea of human adults as indicated by a reduction in DPOAE amplitude. This effect is presumed to be due to activation of medial efferent fibers coursing from mid-brainstem level to the base of OHCs. Consistent with previous experiments, these results suggest that the MOC system can modify mechanical functioning of the cochlea. The mechanism by which this is achieved is less clear. Some evidence suggests that stimulation of the olivocochlear bundle reduces the endocochlear potential that is so critical to OHC function and the generation of DPOAEs (Mills *et al.*, 1993; Mountain, 1980). Because fibers of the medial olivocochlear system predominantly innervate outer-hair cells, it is presumed that they exert their suppressive influ-

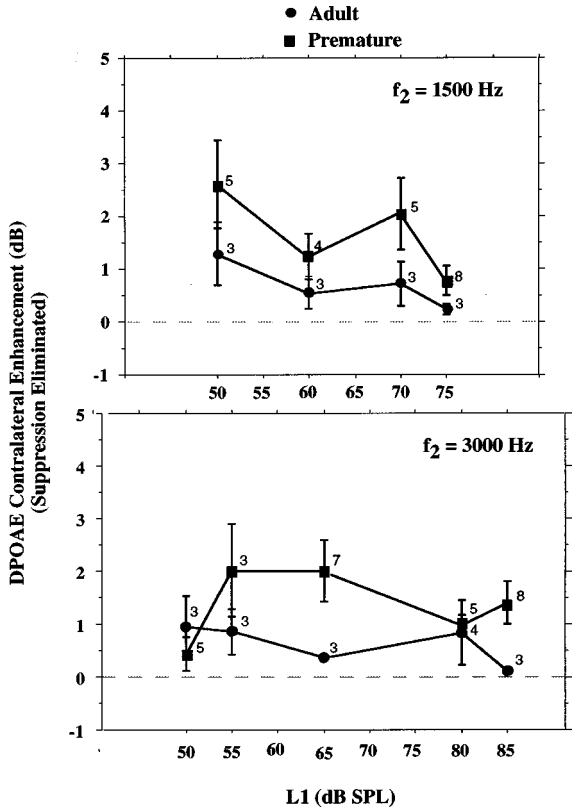


FIG. 6. Mean enhancement is plotted as a function of stimulus levels for $f2 = 1500$ and 3000 Hz. DPOAE contralateral enhancement was calculated by subtracting DPOAE amplitude −BBN from amplitude recorded +BBN after all values reflecting amplitude suppression were eliminated. Data from the two age extremes (adults and premature neonates) are displayed, and only levels with three or more observations are included. Numbers of ears contributing to each average are presented to the right of the symbols. Vertical bars represent ± one standard error of the mean.

ence via this pathway and most probably by interfering with cochlear amplifier function. Activation of the olivocochlear bundle, for example, is thought to modulate OHC motility and consequently, influence cochlear amplifier efficiency and stability (Maison *et al.*, 1997).

DPOAE contralateral suppression has been described in previous studies with human adults (Williams and Brown, 1995, 1997; Moulin *et al.*, 1993). Contralateral suppression values were comparable among these studies, and averaged between 1 to 2 dB. Previously described frequency and level trends related to DPOAE contralateral suppression were also observed in this study. DPOAE contralateral suppression is more pronounced at low stimulus levels and only present at low-to-mid frequencies. Although the level effect appeared more elusive in this study, it was clearly present in the analyses of contralateral suppression with enhancement data eliminated (Fig. 5). The strong effect of medial efferent system fibers at low levels is consistent with an efferent system that has regulatory influence over level-dependent cochlear amplifier activity. Because the cochlear amplifier functions only at low stimulus levels, it is expected that efferent modulation of this process would also occur only at these operative levels.

It is not clear why the medial efferent effect is only present at low-to-mid frequencies in humans, since morphological study of the human cochlea has shown that the density of efferent innervation tends to decrease from base to apex (Nadol, 1990). The robust low-to-mid-frequency efferent response might be related to the spectrum of speech. One of the dominant hypotheses of efferent-system significance is that the olivocochlear pathway facilitates transmission of salient signals (i.e., speech) to the brain for interpretation, while inhibiting nonrelevant stimuli such as background noise (Liberman, 1988). The frequencies most critical to average speech perception and understanding are between 1000 and 4000 Hz (Pavolic, 1987); if the medial efferent system is facilitating transmission of energy in these frequency ranges as hypothesized, we would expect to see the largest contralateral suppressive effect at these frequencies. This is consistent with the results of this study reporting robust suppression at 1500 and 3000 Hz only.

Another hypothesis of efferent system function is that it serves a protective role against acoustic trauma (Patuzzi and Thompson, 1991; Reiter and Liberman, 1995; Zheng *et al.*, 1997). This is based on the inhibitory nature of the efferent effect and the potential application of this inhibition in suppressing damaging acoustic signals. Again, it is probable that the frequency range of maximum protection is the frequency range where communication signals are found. Thus, the frequency-dependent observation of contralateral suppression reported in this study is consistent with theories of efferent system involvement in facilitating speech perception and protection against acoustic trauma.

It is unclear at present what the absence or presence of medial efferent system function (as revealed by DPOAE contralateral suppression) indicates about status of auditory coding and processing in human adults. Our results suggest that even normal-hearing adults with apparently normal auditory functioning can show absence of DPOAE contralateral suppression. Without further testing to differentiate between individuals with and without contralateral suppression, it may be difficult to interpret these findings. For example, it is possible that normal-hearing adults with absent or weak contralateral suppression also have difficulty in speech perception when their auditory system is stressed with competing background noise or degraded in any other way. Recent research has suggested a connection between auditory perception under noisy conditions and function of the efferent system (Liberman, 1988; Micheyl and Collet, 1996). Perhaps adults with absent contralateral suppression in this study could be distinguished from their peers with robust contralateral suppression by using tests of central auditory processing such as speech perception in noise, filtered speech, or dichotic listening tasks. It is not clear what significance absent contralateral suppression has for auditory perception and processing in human adults. This area also needs further exploration.

B. Medial efferent system effects in human neonates

Term neonates show a predominantly suppressive effect when BBN is presented contralaterally. The magnitude and pattern of contralateral suppression in term-born neonates is comparable to adults, suggesting that the medial efferent effect on cochlear function is mature by 40 weeks gestation.

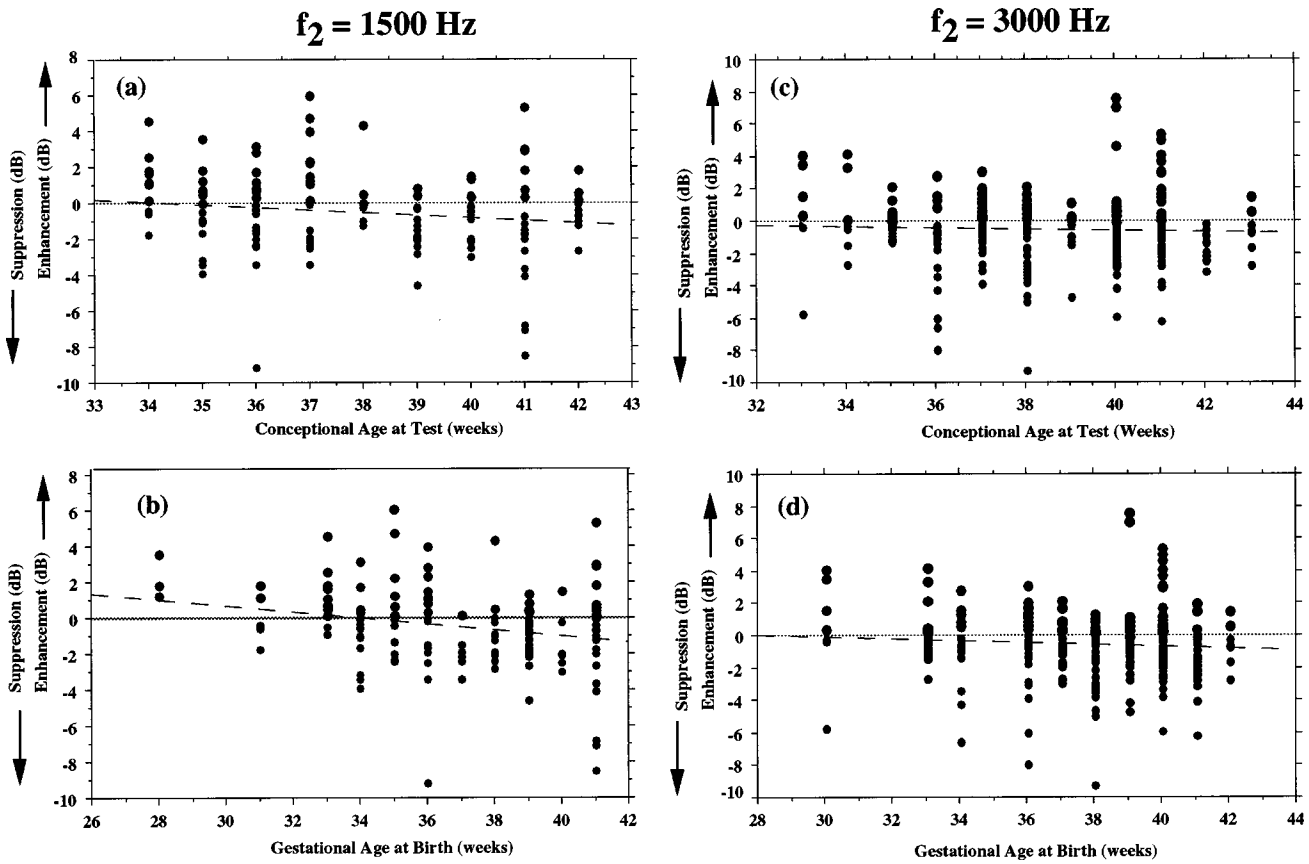


FIG. 7. Scatterplots of contralateral enhancement (points above the horizontal line at 0 dB) and suppression (points below the horizontal line at 0 dB) as a function of conceptional and gestational age. Conceptional age reflects maturational status at the time of evaluation. Gestational age reflects the extent of prematurity at birth. Data for $f2 = 1500$ Hz are on the left side of the figure; data for $f2 = 3000$ Hz are on the right. A regression line is also plotted as a dashed line in each graph.

Results from at least two studies of TEOAE contralateral suppression in neonates support this contention (Ryan and Piron, 1994; Goforth *et al.*, 1997). In contrast, Morlet and colleagues (1993) did not observe significant suppression in neonates ranging from 33 to 39 weeks conceptional age. However, this study involved collection of only one noise and one no-noise condition per subject. Relying on only one observation allows for interference from spurious noise that may impact DPOAE amplitude. In addition, TEOAE contralateral suppression was determined from reproducibility scores including the entire response. This may be problematic since TEOAE suppression is larger and more easily detected after the initial 8–10 ms and often optimally detected within restricted time periods between 8 and 18 ms (Berlin *et al.*, 1994; Hood *et al.*, 1996; Veuillet *et al.*, 1991). By using this analysis strategy, Morlet and colleagues may have minimized their chance of detecting the effect.

TABLE I. Standard error of the mean for DPOAE amplitude and contralateral suppression difference scores at three $f2$ frequencies. Standard error (versus standard deviation) was applied as a measure of intersubject variability because age groups had varying numbers of subjects.

		$f2$ Frequency (Hz)		
		1500	3000	6000
DPOAE amplitude				
Adults	s.e.	1.39	1.33	2.98
Term		1.49	1.03	2.06
Premature		1.44	1.60	2.58
DPOAE contralateral suppression				
Adults	s.e.	0.31	0.34	···
Term		0.50	0.41	···
Premature		0.45	0.55	···

Premature neonates as a group did not show adult-like effects of contralateral broadband noise on DPOAE amplitude. One could argue that it may be more difficult to detect efferent effects in such a heterogeneous group. Premature subjects varied in conceptional age from 33 to 37 weeks at test. The most rapid maturation of human brainstem structure, such as myelin, and function, such as axonal conduction and synaptic transmission, occurs between 28 weeks and term birth (Ponton *et al.*, 1996). Thus, it is evident that premature neonates tested here could have been at different maturational stages although they were only weeks apart in conceptional age. The only way to reduce this heterogeneity is to study contralateral suppression in groups of premature neonates with much smaller age ranges.

In addition to the variability imposed by varying conceptional ages at test, gestational age at birth varied from 27 to 36 weeks in premature subjects. A neonate born at 27 weeks of gestation might suffer a greater disruption of normal sensory and/or neural maturation than a neonate born at 36 weeks, although both neonates are considered premature. Data in Fig. 7(b) support this idea. In general, these data

suggest that the earlier a baby was born, the more likely it is they will show nonadult-like expressions of efferent function (i.e., contralateral enhancement). The extent of prematurity at birth apparently influences this medial efferent system function more than maturational status (i.e., weeks postconception), indicating that early birth is, in and of itself, disruptive to formation of normal efferent system function.

Recent findings in the cat suggest that the developmental period during which efferent fibers are severed is critical in determining the impact this disruption might have on auditory perception and coding (Walsh *et al.*, 1998). Eliminating MOC fibers in adult cats, for example, does not alter the discharge rate or frequency selectivity of VIIIth-nerve fibers. However, when the same procedure is conducted in neonatal kittens prior to appropriate efferent fiber–OHC innervation, VIIIth-nerve tuning is abnormally broad, even in adulthood.

There are many additional birth-related factors that make the premature group heterogeneous and may contribute to efferent system immaturity, such as birth weight, APGAR scores, supplemental oxygen, and duration of extrauterine life. However, there are insufficient preterm subjects in this study to conduct a multiple regression analysis on these additional factors that may impact functional status of the medial efferent system. With a larger sample, such an analysis could be helpful in determining whether any of these factors appropriately categorize the premature neonate with and without DPOAE contralateral suppression.

One could argue that the heterogeneity of premature subjects produced increased variability in their data and may be responsible for the almost equal observation of DPOAE amplitude enhancement (43%) and suppression (57%) in the premature age group. If so, enhancement does not represent a physiologically significant effect and a sign of immaturity as hypothesized. However, standard errors (s.e.) of mean DPOAE amplitude and contralateral suppression from the three age groups are shown in Table I. Premature neonates had larger values of error than adults in some cases, but most often showed error values equal to or smaller than term neonates. Term subjects did show adult-like DPOAE changes with BBN. Thus, it is difficult to attribute the enhancement observed in premature neonates to excessive variability or ''noise'' in DPOAE recordings.

Another potential source of age differences in DPOAE measurement might be immaturity of the outer and middle ear of neonates. Keefe and colleagues have reported poor transfer of energy into the neonatal middle ear and reduced middle-ear compliance in human infants (Keefe *et al.*, 1993, 1994). Thus, it is possible that attenuated primary-tone levels reached the neonatal cochlea, and consequently evoked reduced cochlear output (i.e., low DPOAE amplitude). Because the contralateral noise paradigm records relative rather than absolute DPOAE amplitude (i.e., a difference score), the absolute amplitude of the initial DPOAE (-BBN) is not an important variable and should not affect the contralateral suppression results. Additionally, previous work has clearly indicated that neonatal DPOAEs are not low in amplitude as would be expected if primary tones were greatly attenuated in the middle ear. Rather, DPOAEs are significantly larger in neonates than adults (Abdala, 1996; Lasky, 1998; Prieve *et al.*, 1997).

The most parsimonious explanation for these results is that preterm human neonates do not have mature function of the medial efferent system. This explanation is consistent with human anatomical data. Maturation of the MOC–OHC synapse is among the last event to occur in peripheral auditory development. It is not complete until sometime in the third trimester of pregnancy (Lavigne-Rebillard and Pujol, 1987, 1988; Pujol, 1985), although olivocochlear neurons are identifiable in the human brainstem as early as 21–22 fetal weeks (Moore and Simmons, 1998). Therefore, the youngest infants in this sample (33 weeks conceptional age) could have had nonadult-like changes in DPOAE amplitude due to immature innervation of OHCs by medial efferent fibers.

C. DPOAE amplitude enhancement

Figure 5 suggests that when the enhancement component is removed from DPOAE contralateral suppression data for all subjects, suppression is comparable in magnitude among ages. The enhancement aspect of the efferent effect, not the suppressive aspect, is responsible for the age results observed. The medial olivocochlear system is functional in these premature subjects because DPOAE amplitude is influenced by contralateral noise, but it is simply not adult-like in its expression. For example, DPOAE contralateral enhancement is much more frequent in premature subjects than adults, suggesting that activation of the MOC pathway does affect outer hair-cell output in these neonates, but its effect differs from the effect evoked in the mature system. In addition to frequency of occurrence, the size of the enhancement effect, when present, is greater in premature neonates ($\bar{x} = 1.7$ dB) than adults ($\bar{x} = 0.7$ dB).

Enhancement of DPOAEs has been shown previously in mature animals when certain parameters are varied during test. For example, Siegel and Kim (1982) observed both suppression and enhancement of DPOAEs in chinchilla when electrical shocks were delivered to the crossed olivocochlear bundle. However, enhancement was only observed for certain $f1$ and $f2$ frequency combinations. Liberman *et al.* (1996) observed enhancement of the DPOAE with contralateral noise in cats as the $L1$-$L2$ separation approached 20 dB. In the present experiment, stimulus level separation was held constant and $f1,f2$ frequencies were applied equally to adults and neonates. Absolute level of the primary tones may be different in adult and neonatal ear canals due to the smaller external ear canals in neonates, but the separation of $L1$ and $L2$ is not likely to be differentially affected in adults and neonates. Therefore, the age effect observed (i.e., greater occurrence of enhancement in premature neonates), cannot be easily attributed to the choice of stimulus parameters.

Contralaterally evoked amplitude enhancement may suggest that the efferent system of premature neonates shows an excitatory, rather than a typical, inhibitory response. It is possible that during the initial stages of MOC–OHC synaptic maturation, the effect of medial efferent fibers on OHCs is excitatory. Jenkins and colleagues (1993) recently studied the effects of contralateral sound on VIIIth-nerve fiber responses in kittens. They measured fiber threshold and spon-

taneous rate in kittens of various ages as broadband noise or tone bursts were presented to the contralateral ear, and observed either response suppression, enhancement, or no effect. Neural responses from kittens younger than 10 postnatal days did not change with any form of contralateral stimulation. After 10 days, spontaneous rate enhancement was observed in VIIIth-nerve fibers during noise or tone burst contralateral stimulation. The enhancement effect was very small, as observed in this study, but consistently present in kittens within this age range. The percentage of neural fibers showing rate enhancement decreased with postnatal age, and those showing rate suppression increased.

Jenkins *et al.* hypothesized that the enhancement effect is a transient response reflecting a stage of auditory immaturity. Further, they speculated that enhancement might result from an initially excitatory efferent–OHC synapse. The transition from excitation to inhibition occurs only as the animal's auditory system reaches its mature state. There are various examples within developing neural systems where a transition from excitation to inhibition has been observed (e.g., Wang *et al.*, 1994; Kandler and Friauf, 1995).

In conclusion, the results of this study indicate that the human medial efferent system has a generally adult-like effect on cochlear mechanics by term birth. Results suggest that an immaturity in the medial efferent system may exist just prior to term birth, when an atypically high occurrence of DPOAE amplitude enhancement (vs suppression) is observed in premature neonates when contralateral broadband noise is presented. This phenomenon of DPOAE amplitude enhancement may represent a temporary stage of immaturity in MOC function.

ACKNOWLEDGMENTS

This work was partially supported by a small grant from the American Academy of Audiology and by research grant No. 1 R29 DC03552-01A1 from the National Institutes of Deafness and Other Communication Disorders, National Institutes of Health. Authors would like to thank Sandy Oba for assistance in data collection and management.

Abdala, C., Sininger, Y., and Ma, E. (**1998**). ''Contralateral suppression of the 2 $f1-f2$ DPOAE in human adults and neonates,'' Abstract Assoc. Res. Otolaryngol., St. Petersburg Beach, FL, Vol. 606, p. 152.

Abdala, C. (**1996**). ''DPOAE (2 $f1-f2$) amplitude as a function of $f2/f1$ frequency ratio and primary tone level separation in human adults and neonates,'' J. Acoust. Soc. Am. **100**, 3726–3740.

Berlin, C., Hood, L., Hurley, A., and Wen, H. (**1994**). ''Contralateral suppression of otoacoustic emissions: An index of the function of the medial olivocochlear system,'' Otolaryngol.-Head Neck Surg. **110**, 3–21.

Buno, Jr., W. (**1978**). ''Auditory nerve fiber activity influenced by contralateral ear sound stimulation,'' Exp. Neurol. **59**, 62–74.

Capps, M., and Ades, H. (**1968**). ''Auditory frequency discrimination after transection of the olivocochlear bundle in squirrel monkeys,'' Exp. Neurol. **21**, 147–158.

Carlier, E., and Pujol, R. (**1982**). ''Sectioning the efferent bundle decreases cochlear frequency selectivity,'' Neurosci. Lett. **28**, 101–106.

Collet, L., Kemp, D., Veuillet, E., Duclaux, R., Moulin, A., and Morgon, A. (**1990**). ''Effect of contralateral auditory stimuli on active cochlear micromechanical properties in human subjects,'' Hearing Res. **43**, 251–262.

Folsom, R., and Owsley, R. (**1987**). ''$N1$ action potentials in humans. Influence of simultaneous contralateral stimulation,'' Acta Oto-Laryngol. **103**, 262–265.

Galambos, R. (**1956**). ''Suppression of auditory nerve activity by stimulation of efferent fibers to cochlea,'' J. Neurophysiol. **19**, 424–437.

Goforth, L., Hood, L., and Berlin, C. (**1997**). ''Efferent suppression of transient-evoked otoacoustic emissions in human infants,'' Abstract Assoc. Res. Otolaryngol., St. Petersburg Beach, FL, Vol. 662, p. 166.

Guinan, J., and Gifford, M. (**1988**). ''Effects of electrical stimulation of efferent olivocochlear neurons on cat auditory nerve fibers III: Tuning curves and threshold at CF.'' Hearing Res. **37**, 29–46.

Hood, L., Berlin, C., Hurley, A., Cecola, P., and Bell, B. (**1996**). ''Contralateral suppression of transient evoked otoacoustic emissions in humans: Intensity effects,'' Hearing Res. **101**, 113–118.

Igarishi, M., Cranford, J., Nakai, Y., and Alford, B. (**1979**). ''Behavioral auditory function after transection of crossed olivo-cochlear bundle in the cat,'' Acta Oto-Laryngol. **87**, 79–83.

Jenkins, J., McGee, J., and Walsh, E. (**1993**). ''Developmental changes of auditory nerve responses to efferent stimulation,'' Soc. Neurosci Abstr. **23**(1), 534.

Johnstone, B., Patuzzi, R., and Yates, G. (**1986**). ''Basilar membrane measurements and the traveling wave,'' Hearing Res. **22**, 147–153.

Kandler, K., and Friaur, E. (**1995**). ''Development of glycinergic and glutamatergic synaptic transmission in the auditory brainstem of perinatal rats,'' J. Neurosci. **15**, 6890–6904.

Kawase, T., and Liberman, M. (**1993**). ''Antimasking effects of the olivocochlear reflex I: Enhancement of compound action potentials to masked tones,'' J. Neurophysiol. **70**, 2519–2532.

Keefe, D., Bulen, J., Arehart, K., and Burns, E. (**1993**). ''Ear-canal impedance and reflection coefficient in human infants and adults,'' J. Acoust. Soc. Am. **94**, 2617–2638.

Keefe, D., Bulen, J., Arehart, K., and Burns, E. (**1994**). ''Pressure transfer function and absorption cross section from the diffuse field to the human infant ear canal,'' J. Acoust. Soc. Am. **95**, 355–371.

Lasky, R. (**1998**). ''Distortion product otoacoustic emissions in human newborns and adults: II. Level Effects,'' J. Acoust. Soc. Am. **103**, 992–1000.

Lavigne-Rebillard, M., and Pujol, R. (**1987**). ''Surface aspects of the developing human organ of Corti,'' Acta Oto-Laryngol. Suppl. **436**, 43–50.

Lavigne-Rebillard, M., and Pujol, R. (**1988**). ''Hair cell innervation in the fetal human cochlea,'' Acta Oto-Laryngol. **105**, 398–402.

Liberman, M. C., Puria, S., and Guinan, J. (**1996**). ''The ipsilaterally evoked olivocochlear reflex causes rapid adaptation of the 2 $f1-f2$ distortion product otoacoustic emission,'' J. Acoust. Soc. Am. **99**, 3572–3584.

Liberman, M. C. (**1988**). ''Response properties of cochlear efferent neurons: Monaural versus binaural stimulation and the effects of noise,'' J. Neurophysiol. **60**, 1779–1798.

Liberman, M. C. (**1989**). ''Rapid assessment of sound-evoked olivocochlear feedback: Suppression of compound action potential by contralateral sound,'' Hearing Res. **38**, 47–56.

Littman, T., Cullen, J., and Bobbin, R. (**1992**). ''The effect of olivocochlear bundle transection on tuning curves and acoustic distortion products,'' J. Acoust. Soc. Am. **92**, 1945–1952.

Maison, S., Micheyl, C., Chays, A., and Collet, L. (**1997**). ''Medial olivocochlear system stabilizes active cochlear micromechanical properties in humans,'' Hearing Res. **113**, 89–98.

Micheyl, C., and Collet, L. (**1996**). ''Involvement of the olivocochlear bundle in the detection of tones in noise,'' J. Acoust. Soc. Am. **99**, 1604–1610.

Mills, D., Norton, S., and Rubel, E. (**1993**). ''Vulnerability and adaptation of distortion product otoacoustic emissions to endocochlear potential variation,'' J. Acoust. Soc. Am. **94**, 2108–2122.

Moore, J., and Simmons, D. (**1999**). ''The human olivocochlear system: Organization and development,'' Audiol. Neuro-Otol. (in press).

Morlet, T., Collet, L., Salle, B., and Morgon, A. (**1993**). ''Functional maturation of cochlear active mechanisms and of the medial olivocochlear system in humans,'' Acta Oto-Laryngol. **113**, 271–277.

Mott, J., Norton, S., Neely, S., and Warr, W. (**1989**). ''Changes in spontaneous otoacoustic emissions produced by acoustic stimulation of the contralateral ear,'' Hearing Res. **38**, 229–242.

Mountain, D. (**1980**). ''Changes in endolymphatic potential and crossed olivocochlear bundle stimulation alter cochlear mechanics,'' Science **210**, 71–72.

Moulin, A., Collet, L., and Duclaux, R. (**1993**). ''Contralateral auditory stimulation alters acoustic distortion products in humans,'' Hearing Res. **65**, 193–210.

Nadol, J. (**1990**). ''Synaptic morphology of inner and outer hair cells of the human organ of corti,'' J. Electron Microsc. **15**, 187–196.

Patuzzi, R., and Thompson, M. (**1991**). ''Cochlear efferent neurones and protection against acoustic trauma: protection of outer hair cell receptor current and interanimal variability,'' Hearing Res. **54**, 45–58.

Pavolic, C. (**1987**). ''Derivation of primary parameters and procedures for use in speech intelligibility predictions,'' J. Acoust. Soc. Am. **82**, 413–422.

Ponton, C., Moore, J., and Eggermont, J. (**1996**). ''Auditory brain stem response generation by parallel pathways: Differential maturation of axonal conduction time and synaptic transmission,'' Ear Hear. **17**, 402–410.

Prieve, B., Fitzgerald, T., Schulte, L., and Kemp, D. (**1997**). ''Basic characteristics of distortion product otoacoustic emissions in infants and children,'' J. Acoust. Soc. Am. **102**, 2871–2879.

Puel, J., and Rebillard, G. (**1990**). ''Effect of contralateral sound stimulation on the distortion product $2f1-f2$: Evidence that the medial efferent system is involved,'' J. Acoust. Soc. Am. **87**, 1630–1635.

Pujol, R. (**1985**). ''Morphology, synaptology and electrophysiology of the developing cochlea,'' Acta Oto-Laryngol. **421**, 5–9.

Puria, S., Guinan, J., and Liberman, M. (**1996**). ''Olivocochlear reflex assays: Effects of contralateral sound on compound action potentials versus ear-canal distortion products,'' J. Acoust. Soc. Am. **99**, 500–507.

Rajan, R. (**1990**). ''Electrical stimulation of the inferior colliculus at low rates protects the cochlea from auditory desensitization,'' Brain Res. **506**, 192–204.

Reiter, E., and Liberman, M. C. (**1995**). ''Efferent-mediated protection from acoustic overexposure: relation to slow effects of olivocochlear stimulation,'' J. Neurophysiol. **73**, 506–514.

Ryan, S., and Piron, J. (**1994**). ''Functional maturation of the medial efferent olivocochlear system in human neonates,'' Acta Oto-Laryngol. **114**, 485–489.

Scharf, B., Magnan, J., Collet, L., Ulmer, E., and Chays, A. (**1994**). ''On the role of the olivocochlear bundel in hearing: A case study,'' Hearing Res. **75**, 11–26.

Siegel, J., and Kim, D. (**1982**). ''Efferent neural control of cochlear mechanics? Olivocochlear bundle stimulation affects cochlear biomechanical nonlinearity,'' Hearing Res. **6**, 171–182.

Silman, S. (**1984**). *The Acoustic Reflex: Basic Principles and Clinical Applications* (Academic, Orlando, FL).

Sininger, Y., Abdala, C., and Cone-Wesson, B. (**1997**). ''Auditory threshold sensitivity of the human neonate as measured by the auditory brainstem response,'' Hearing Res. **104**, 27–38.

Veuillet, E., Collet, L., and Duclaux, R. (**1991**). ''Effect of contralateral acoustic stimulation on active cochlear micromechanical properties in human subjects: Dependence on stimulus variables,'' J. Neurophysiol. **65**, 724–735.

Walsh, E., McGee, J., McFadden, S., and Liberman, M. (**1998**). ''Long-term effects of sectioning the olivocochlear bundle in neonatal cats,'' J. Neurosci. **18**, 3859–3869.

Wang, J., Reichling, D., Krozis, A., and MacDermott, A. (**1994**). ''Developmental loss and glycine-induced depolarization and $Ca2+$ transients in embryonic rat dorsal horn neurons in culture,'' Eur. J. Neurosci. **6**, 1275–1280.

Williams, D., and Brown, A. (**1995**). ''Contralateral and ipsilateral suppression of the $2f1-f2$ distortion product in human subjects,'' J. Acoust. Soc. Am. **97**, 1130–1140.

Williams, D., and Brown, A. (**1997**). ''The effect of contralateral broadband noise on acoustic distortion products from the human ear,'' Hearing Res. **104**, 127–146.

Winslow, R., and Sachs, M. (**1987**). ''Effect of electrical stimulation of the crossed olivocochlear bundle on auditory nerve response to tones in noise,'' J. Neurophysiol. **57**, 1002–1021.

Zheng, X.-Y., Henderson, D., Hu, B.-H., Ding, D.-L., and McFadden, S. (**1997**). ''The influence of the cochlear efferent system on chronic acoustic trauma,'' Hearing Res. **107**, 147–159.

Spontaneous otoacoustic emissions in heterosexuals, homosexuals, and bisexuals

Dennis McFadden[a] and Edward G. Pasanen
Department of Psychology and Institute for Neuroscience, Mezes Hall 330, University of Texas, Austin, Texas 78712

(Received 17 September 1998; accepted for publication 18 December 1998)

Click-evoked otoacoustic emissions (CEOAEs) were previously shown to be significantly less strong in homosexual and bisexual females than in heterosexual females. Here it is reported that the spontaneous otoacoustic emissions (SOAEs) of those same 60 homosexual and bisexual females were less numerous and weaker than those in 57 heterosexual females. That is, the SOAEs of the homosexual and bisexual females were intermediate to those of heterosexual females and heterosexual males. The SOAE and CEOAE data both suggest that the cochleas of homosexual and bisexual females have been partially masculinized, possibly as part of some prenatal processes that also masculinized whatever brain structures are responsible for sexual orientation. For males of all sexual orientation, the SOAEs were less numerous and weaker than for the females, and there were no significant differences among the 56 heterosexual, 51 homosexual, and 11 bisexual males. All subjects passed a hearing screening test. When all SOAEs above 3000 Hz were excluded (as a control against incipient, undetected hearing loss) the same results were obtained as with the full range of data (550–9000 Hz). The differential use of oral contraceptives by the heterosexual and nonheterosexual females also could not explain the differences in their OAEs. © *1999 Acoustical Society of America.* [S0001-4966(99)00804-8]

PACS numbers: 43.64.Jb, 43.64.Gz, 43.64.Bt [BLM]

INTRODUCTION

Normal cochleas have the ability to produce sounds as well as to receive and process sounds. These cochlea-generated sounds propagate back through the middle-ear system into the external auditory canal, where they can be recorded using miniature microphone systems. There are several types of these cochlea-generated sounds which are known collectively as otoacoustic emissions or OAEs (for a review, see Probst *et al.*, 1991).

Two types of OAE are of interest here. Click-evoked otoacoustic emissions (CEOAEs) are echolike waveforms that are produced in response to brief acoustic stimuli (Kemp, 1978, 1979). CEOAEs can last tens of milliseconds, and essentially all normal-hearing ears produce them, although their strength varies across ears. Spontaneous otoacoustic emissions (SOAEs) are essentially tonal sounds that are produced continuously by the majority of normal-hearing ears. Both CEOAEs and SOAEs exhibit sex and ear differences (for summaries see Talmadge *et al.*, 1993; McFadden *et al.*, 1996). CEOAEs are stronger and SOAEs are more numerous both in females than males, and in right ears than left (McFadden, 1998). These same patterns of sex and ear differences exist in infants and children (Burns *et al.*, 1992; Norton, 1992), and the available longitudinal data indicate that OAEs remain reasonably constant through life (Kemp *et al.*, 1986; Franklin *et al.*, 1992; Burns *et al.*, 1993, 1994; Prieve *et al.*, 1993). Accordingly, it seems safe to think of OAEs as being reasonably stable traits through life.

In order to explain various facts of cochlear function in mammals, an active mechanism called the cochlear amplifier has been proposed (Davis, 1983). The contribution of the individual cochlear amplifiers hypothesized to be arrayed along the length of the basilar membrane is to magnify membrane displacement at low sound-pressure levels, thereby giving the auditory system its high sensitivity to weak sounds. Mammalian OAEs are widely regarded to be unintended by-products of the normal functioning of these hypothesized cochlear amplifiers. Thus, both auditory sensitivity and OAEs are regarded to be consequences of a common mechanism, the cochlear amplifiers. According to common belief, the amplifiers cannot perform optimally unless the relatively fragile outer hair cells are intact.

We recently reported that CEOAEs are weaker in homosexual and bisexual females than in heterosexual females (McFadden and Pasanen, 1998). That is, the strength of the CEOAEs in nonheterosexual females was intermediate to that in heterosexual females and heterosexual males. By inference, then, the cochlear amplifiers of nonheterosexual females can be presumed to be less strong than those of heterosexual females, and to be intermediate to those of heterosexual females and males. Here we report that the SOAEs of those same subjects showed the same basic pattern as did their CEOAEs. The SOAEs of homosexual and bisexual females were less numerous and less strong than those of heterosexual females, and were intermediate to those of heterosexual females and heterosexual males. This additional evidence of functional differences in the cochleas of nonheterosexual females bolsters the interpretation that their peripheral auditory systems have been masculinized, possibly at the same stage of development when whatever

[a]Electronic mail: mcfadden@psy.utexas.edu

brain structures are responsible for sexual orientation also were masculinized.

I. METHODS

The SOAE measurements to be reported came from the same subjects as in the CEOAE paper (McFadden and Pasanen, 1998). A description of the recruiting procedures, the average age in the various subject groups, the number of subjects discarded for hearing loss and other reasons, and the methods for determining sexual orientation can be found in that previous paper.

For consistency, all subjects discarded from the previous analyses because of hearing loss (17 females, 24 males) or questionable CEOAE data (6 females, 7 males) were also discarded here. In addition, the SOAE data from one homosexual female and one heterosexual male were omitted from all analyses here because of evidence that the gain of an audio amplifier was set incorrectly for the SOAE measurements in one ear. This left 235 subjects out of 291 recruited to the experiment. Subjects were classified into six categories of sexual orientation on the basis of their responses to a series of items on a questionnaire; the number of subjects of each type is indicated in Fig. 1. For some of the within-subjects analyses reported below, only subjects having at least one SOAE in each ear could be included. That reduced the overall number of subjects to 132, with the individual *N*'s being 46, 24, and 16 for the female heterosexuals, homosexuals, and bisexuals, respectively, and 26, 17, and 3 for the male heterosexuals, homosexuals, and bisexuals, respectively.

The collection of the SOAE data was interleaved with the collection of the CEOAE data. For the odd-numbered subjects in each group, the order of data collection was SOAEs right ear, CEOAEs right ear, CEOAEs left ear, SOAEs left ear, and the order was reversed for the even-numbered subjects. For the odd-numbered subjects, OAEs were measured before the subject completed a lengthy questionnaire, and that order was reversed for the even-numbered subjects. An experimental session took approximately two hours. Subjects were paid $30 for their time.

For the OAE measurements, the subject lay on a small cot in a darkened, sound-proofed room. The foam-covered tip of an Etymotic ER-10A insert microphone system was inserted into the external ear canal. The probe tip contained two stainless steel tubes, one leading (via a plastic tube) to the earphone used to present the clicks used for eliciting CEOAEs. The microphone output passed through an ER10-72 pre-amplifier and then through a custom-built, low-noise amplifier/filter that amplified the signal by about 30 dB and high-passed it above about 400 Hz (in order to remove extraneous body noises). The output of the amplifier/filter was delivered to both a spectrum analyzer (Hewlett-Packard model 35665A) and an analog-to-digital converter (National Instruments A2100) located in a Macintosh Quadra 950 computer. In an attempt to avoid initializing effects (Whitehead, 1991; McFadden and Pasanen, 1994; Smurzynski and Probst, 1998), the subject lay quietly in the test room with the probe tip in place for at least 15 min prior to collection of the first SOAE data.

The SOAE measurements for each ear were extracted from four 30-s samples of the sounds monitored by the miniature microphone in the external ear canal. These waveforms were digitized (with 16-bit resolution) at a sampling rate of 22 050 sample points per second and stored for later analysis. Prior to the collection of each time sample, the subject was informed over an intercom that a data sample was about to be collected and that he or she should lie as quietly as possible until told that the 30-s period was over. Upon collection of the first 30-s sample, a brief analysis was performed to verify overall system performance as well as the subject's ability to lie quietly. Subsequent samples were also analyzed when it was deemed necessary. Accordingly, successive 30-s samples were often separated by several tens of seconds. Some especially noisy subjects were reinstructed about the necessity of lying still and additional 30-s samples were collected as substitutes for noisy ones. While these time samples were being collected, spectra were also being collected and displayed on the spectrum analyzer as general information for the experimenters, to verify the integrity of the monitoring system, and to monitor the noise level of the subject.

The identification and measurement of SOAEs from the obtained time waveforms was accomplished using a fully objective, computerized procedure. An earlier version has been described (Pasanen and McFadden, 1996), and a report on the final version is being prepared. There are two aspects to the extraction of SOAEs: distinguishing the spectral peaks that are likely to correspond to true SOAE sources from the random fluctuations in the spectrum, and then assigning estimates of frequency and magnitude to those peaks. For experiments like the present one, involving large numbers of subjects and questions about the relative prevalence of SOAEs in different groups, it is clearly desirable to have an automated, objective procedure for identifying the SOAEs. The procedure used here had several stages that will be described only briefly because of space limitations.

The first goal was to eliminate the noisiest time segments from further analysis. To accomplish this, the rms power was calculated for successive 743-ms segments in the 2 min of data collected for each ear. For these calculations, a new segment began every 185.8 ms (corresponding to 75% overlap in the segments). Then, the 150 segments having the lowest levels were selected and a 16 384-point FFT (fast Fourier transform) calculated for each. For a 2-min data sample, this criterion was roughly equivalent to using the quietest 25% of all segments. The resulting power spectra were summed to yield what shall be called the initial power spectrum for that ear. Each spectral point in the initial spectrum represented the acoustic power in a bin approximately 1.35 Hz in width. Bins below 300 Hz and above 9500 Hz were eliminated from further analyses.

The next goal was to obtain estimates of what the power spectrum would have been had there been no SOAEs; that is, to obtain estimates of the noise floor in the vicinity of all potential SOAEs. The first step in this process was to delete from the initial spectrum all frequency bins whose power deviated significantly from that in an adjacent region. This was accomplished by stepping through the initial spectrum

(from low frequencies to high) one bin at a time, and, for each bin, comparing its value with the mean and standard deviation of an adjacent range of 30 frequency bins, and deleting that bin value if it exceeded 3 standard deviations (s.d.s) above that running mean. The next step was to replace the deleted values with estimates of what the noise floor would have been had there been no peak. This was accomplished by fitting a straight line to the same range of adjacent frequency bins described above, and using that line to extrapolate a value for each deleted value. The range of adjacent frequency bins used for extrapolation was fixed at 30 bins, sufficiently large to cover a number of cycles of the fine-grain peaks and valleys of the initial spectrum. (The intent was for the fitted line to capture both the average spectral level and the general trend of spectral values in that local region.) The nearer edge of this extrapolation interval was displaced in frequency from the point to be estimated by a number of bins that was directly proportional to the frequency of the estimated point. The displacement was 9 bins (12.15 Hz) at 900 Hz, and 90 bins at 9000 Hz. (Factors such as the minimal spacing and typical widths of true SOAEs affected the choices for the range of the extrapolation interval and its displacement from the deleted bin being estimated.) Following this substitution process, the entire resulting spectrum was smoothed using a simple 51-bin boxcar window. After the boxcar smoothing, those points in the power spectrum that had resulted from extrapolation were again deleted, and the extrapolation process was repeated for them (working from high frequencies to low), followed by another boxcar smoothing. The result will be called the smoothed spectrum. This procedure produced noise floors that were within tenths of a decibel of each other for all six subject groups at each of several spectral regions examined; that is, the smoothed spectra were not systematically different across subject groups.

The next goal was to identify the likely SOAEs. This was accomplished in three steps. First, an estimate of the inherent variability of the spectrum was obtained. Regions of the spectrum in which no extrapolation and substitution were done were scanned for all positive peak deflections, where a peak deflection was defined simply as any bin where the local slope of the spectrum changed from positive to negative. In most cases this was a bin whose value was greater than those of the two flanking bins. This scan typically netted about 1500 peak deflections. The mean of the five spectral bins centered at each such peak deflection was computed, both for the initial spectrum and the smoothed spectrum, and the difference between the two values was calculated and called a ''peak deviation.'' A positive peak deviation indicated a spectral deflection above the smoothed spectrum. The mean and standard deviation of this set of 1500 peak deviations were computed for each ear. (Note that this set of values comprises the distribution of those spectral peaks not likely to be identified as SOAEs.) The next step was to express as a deviation score *every* peak in the initial spectrum, and to compare it with the distribution of inherent variability just obtained. This was accomplished by expressing the five-bin deviation score for a given peak as a ''z-score,'' by taking the difference between it and the mean of the distribution above, and dividing by the standard deviation. When the result was greater than 5.0 s.d. units, that peak deflection was tentatively identified as an SOAE. After all tentative SOAE peaks were identified, another pass was made, to identify any candidate SOAE peaks within 0.1 octave of each other, in accord with the general belief that two true SOAEs cannot be this close in frequency (Zwicker, 1990). In such a case, the tentative SOAE peak having the highest z-score was identified as an SOAE, and any candidate peaks within 0.1 octave of it were discarded. No attempt was made to systematically eliminate those identified SOAEs that might have been distortion products generated by the interaction of other SOAEs (see Probst *et al.*, 1991), although some of these may have been eliminated inadvertently by the 0.1-octave rule. The frequency assigned to an identified SOAE was simply the frequency bin having the highest peak deflection. Only identified SOAEs between 550 Hz and 9000 Hz were included here.

The final goal was to assign an appropriate magnitude to the identified SOAEs. Because all SOAEs contained power in more than one 1.35-Hz frequency bin, using the power in the single bin containing the SOAE peak would have underestimated the total power of the SOAE, sometimes greatly. It was assumed that, in the initial spectrum, the observed power in any frequency bin in the vicinity of an SOAE was a sum of two independent quantities: the residual noise background and power from the SOAE. The smoothed spectrum was used as an estimate of the noise background. To estimate the contribution of the SOAE to the initial spectrum, the smoothed spectrum was subtracted from the initial spectrum, bin by bin (using units of power). Then, beginning at each identified SOAE peak and working upward in frequency, a running sum of bin values was computed, with successively more remote flanking bins added to the sum, provided that each additional bin incremented the running sum by at least 1%. When this criterion was not satisfied, the summing stopped, and the process was repeated, working downward in frequency. The result was an estimate of the total power of the SOAE. That estimate was then summed with the corresponding estimates for all of the other identified SOAEs in that ear, if any, and the result was converted into decibels sound-pressure level (SPL) *re*: 20 μPa and used as the estimate of overall SOAE power for that ear (or first divided by the number of contributing SOAEs and then converted into decibels to obtain an estimate of power per SOAE).[1]

The values of the various parameters used in this objective SOAE-detection algorithm were all selected on the basis of considerable experience with the spectra collected from a number of subjects having varying numbers of SOAEs of differing strength. The result is judged to be a good compromise that is conservative as to what is identified as an SOAE. In general, the weakest peaks identified as SOAEs had five-bin peak deviations of about 1.3 dB or single-bin peak values of about 1.5 dB above the local smoothed spectrum. When the detection criterion, or some other parameter of the algorithm, was changed, the magnitude of the weakest peak characterized as an SOAE sometimes changed considerably, but there was essentially no change in the substance of the results to be reported here.

II. RESULTS

Variables of three general types were of interest here: the number of SOAEs exhibited by an ear (or a person), the strength of those SOAEs, and the proportion of subjects in each group exhibiting SOAEs. Of these, the number of SOAEs exhibited was of greatest interest, in part because all 235 subjects could be included in the analyses. Most of the analyses to be reported were concerned with the full range of identified SOAEs from 550 Hz to 9000 Hz.

A. Number of SOAEs

Figure 1 shows the average number of SOAEs per ear for the left and right ears of the subjects in the six categories of sexual orientation. Two prominent features of these data have been reported previously. The greater number of SOAEs in females than males is a well-documented effect (among subjects previously unselected for sexual orientation), as is the greater number of SOAEs in right ears than left ears (e.g., Talmadge *et al.*, 1993; Penner *et al.*, 1993; McFadden and Loehlin, 1995).

The new findings revealed by the data of Fig. 1 are the obvious differences in the number of SOAEs emitted by people of different sexual orientation. Specifically, homosexual and bisexual females clearly had fewer SOAEs than did the heterosexual females, while there was no apparent difference between heterosexual and homosexual males. This general pattern of results parallels that in the CEOAE data (McFadden and Pasanen, 1998). In both cases, the data for the nonheterosexual females were about halfway between those of the heterosexual females and males. The primary difference between the SOAE and CEOAE data is in the results for the bisexual males which here are generally in line with those from the other males, whereas the CEOAEs of these same subjects appeared to be not fully masculinized. As can be seen, however, the number of bisexual male subjects was quite small, in accord with past reports (see Pattatucci and Hamer, 1995). (Bisexuals were not actively recruited to the experiment but were identified after the fact on the basis of their questionnaire responses.)

For the purpose of statistical analysis, the number of SOAEs in each ear (X) was log transformed as $\ln(X+1)$ in an attempt to make the underlying distributions less skewed. The skewness of every transformed distribution was smaller than that of the untransformed distribution by a factor ranging from 1.8 to 33, and kurtosis was also comparably reduced in the majority of cases. Preliminary analyses of the transformed data revealed no significant difference between the homosexual and bisexual females or between the homosexual and bisexual males. Accordingly, the homosexuals and bisexuals were combined within sex for the main analysis, which used a two-factor ANOVA. The factors were subject type (four levels: heterosexual males and females and nonheterosexual males and females) X ear of test (two levels), with repeated measures on the latter. At the right in Table I are the summary statistics for the four groups created by pooling the homosexuals and bisexuals, and at the left are the results of the ANOVA (on the log transforms of the number of SOAEs). The two main effects were both significant and their interaction was not (first three rows at left). Means comparisons (remaining rows at left of Table I) revealed that heterosexual females were significantly different from homosexual and bisexual females combined, and the large difference between heterosexual females and heterosexual males was also significant. By comparison, heterosexual males were not significantly different from homosexual and bisexual males combined. When the same analysis was performed on the untransformed numbers of SOAEs, all the statistical decisions were exactly the same, which is in accord with the well-known robustness of ANOVA to deviations from normality (Lindman, 1974). [All ANOVAs were done using SuperANOVA (Abacus Concepts, Berkeley, CA, 1989).]

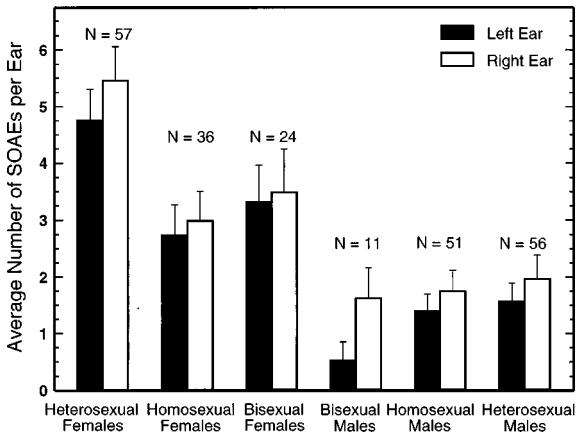


FIG. 1. Average number of SOAEs per ear for each of the six categories of sexual orientation, over the frequency range 550–9000 Hz. The values of N indicate the number of subjects contributing to each mean. The error flags indicate one standard error of the mean.

B. Power of the SOAEs

In the top panel of Fig. 2, the average overall power of the SOAEs is shown for the six categories of sexual orientation. This measure was obtained by summing the powers of all the SOAEs in an ear, expressing the result in decibels sound-pressure level, and averaging across all emitting ears in each subject group.[1] Because the measure overall SOAE power has no meaning for any ear having no SOAEs, all ears having zero SOAEs were discarded for the presentation in Fig. 2; hence the different N's for the two ears in each category.

Clearly, the pattern of results across groups in Fig. 2 appears quite similar to that for the average number of SOAEs in Fig. 1, and for CEOAE amplitude (McFadden and Pasanen, 1998). After preliminary analyses indicated no differences between the homosexual and bisexual females or between the homosexual and bisexual males, those groups were again pooled, and the same $4X2$ ANOVA as used above was applied to the resulting data. Because there is no meaning to any measure of SOAE power when an ear has no SOAEs, and because this is a repeated-measures ANOVA, only subjects having SOAEs in both ears could be included in this analysis, and that decreased the number of subjects to 132. As the results at the top middle of Table II reveal, the two main effects were again significant and their interaction

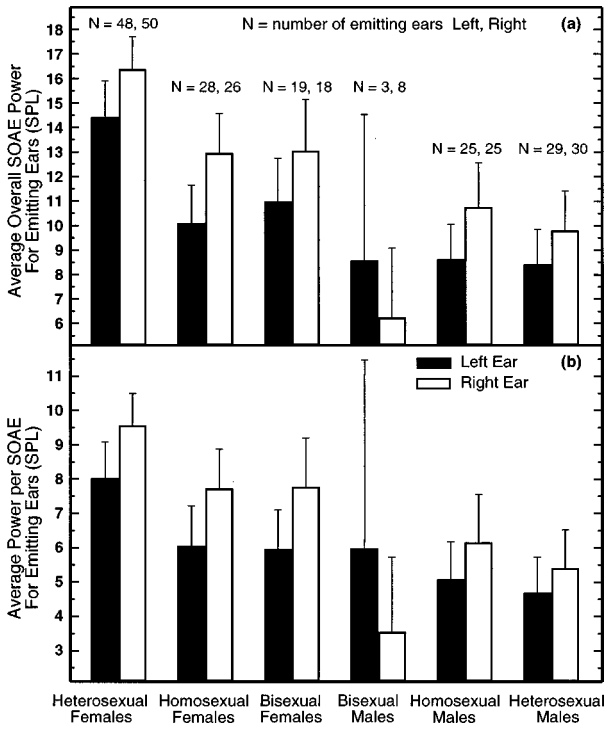


FIG. 2. (a) The average overall SOAE power for emitting ears for each of the six categories of sexual orientation. For each ear, the powers were summed across all SOAEs over the frequency range 550–9000 Hz, and the sum was converted to decibels sound-pressure level. Because ears with no SOAEs could not contribute to this measure, the *N*'s are different for the two ears (but are the same for both panels of this figure). The error flags indicate one standard error of the mean. (b) Same as above except the measure plotted is the average power per SOAE for emitting ears. For each ear, total power across SOAEs was divided by the number of SOAEs contributing to that total, and the result was converted to decibels sound-pressure level.

was not. Means comparisons showed that, with these reduced sample sizes, the large difference between heterosexual females and heterosexual males was still significant, but heterosexual females were only marginally different from nonheterosexual females. Heterosexual males again were not significantly different from nonheterosexual males. (For purposes of comparison, Table II also contains at the left the results of ANOVAs conducted on the measure number of SOAEs using only the 132 subjects included in the analyses of SOAE power; also compare to Table I.)

Logically, the diminished overall SOAE power in homosexual and bisexual females (Fig. 2, top panel) could simply be the consequence of their having fewer SOAEs (Fig. 1). As a check on this possibility, an additional analysis of SOAE magnitude was performed. Instead of simply summing the total powers of all the SOAEs in an ear and expressing the result in decibels (as was done above), the overall power in each ear was first divided by the number of SOAEs contributing to that sum, and the quotient expressed in decibels. That is, what was calculated was the average power per SOAE in an emitting ear instead of the overall power across all SOAEs in that ear. The results are shown in the bottom panel of Fig. 2. Clearly, the pattern of the data across groups is exactly the same as in the top panel of Fig. 2, but the differences were smaller, so when the same two-factor ANOVA was applied to these data (at the right in Table II), the main effects were only marginally significant, and only the means comparison between heterosexual females and heterosexual males was significant. The best interpretation of this outcome becomes more clear after the following analysis.

C. Issue of hearing loss

Even though our subjects were screened for normal hearing, it is still possible that the differences described above between heterosexual and nonheterosexual females were the result of the nonheterosexuals having, for some reason, more (subclinical) hearing loss than heterosexual females. Noise-induced hearing loss and various types of drug-induced hearing loss do diminish the strength of the cochlear amplifiers, and thus diminish OAEs (e.g., Norton *et al.*, 1989; McFadden and Plattsmier, 1984; Long and Tubis, 1988; McFadden and Pasanen, 1994). Most forms of acquired hearing loss affect the highest frequencies first and gradually migrate toward lower frequencies, and visual examination of histograms of the number of SOAEs across acoustic frequency did suggest that homosexual and bisexual females had fewer high-frequency SOAEs than heterosexual females. A strategy for testing whether differences in high-frequency hearing loss are responsible for the parallel group differences seen in Figs. 1 and 2 is to ignore the SOAEs from the highest-frequency regions. Accordingly, the SOAE data were reanalyzed using only the SOAEs between 550 and 3000 Hz instead of the 550–9000 Hz range used to this point. The most important finding from this reanalysis was that the patterns of the data across subject groups were exactly the same as shown in Figs. 1 and 2 (which is why these data are not shown), as were most of the conclusions based on the statistical tests. The bottom half of Table II reveals that, in fact, the differences between the heterosexual and nonheterosexual females were even stronger in this reanalysis than when the full frequency range of SOAEs was included, and some comparisons achieved significance that did not in the previous analyses. This outcome was confirmed by also reanalyzing the measure number of SOAEs using all 235 subjects and only the SOAEs below 3000 Hz (compare to Table I). Both main effects were statistically significant and their interaction was not. For subject type, $F(3,231) = 15.5$, $p = 0.0001$, and for ear difference, $F(1,231) = 11.0$, $p = 0.001$. The difference between heterosexual females and heterosexual males was significant, $F(1,231) = 30.8$, $p = 0.0001$, as was the difference between heterosexual and nonheterosexual females, $F(1,231) = 7.3$, $p = 0.007$, but the difference between heterosexual and nonheterosexual males was not, $F(1,231) = 0.08$, $p = 0.77$.

The conclusion that emerges from these reanalyses is that a simple high-frequency hearing loss is not a likely explanation for the group differences seen in Figs. 1 and 2, or in the CEOAE data (McFadden and Pasanen, 1998). This finding is also relevant for the differences between heterosexual females and heterosexual males because one long-standing explanation for the difference in hearing sensitivity (and, by implication, in OAEs) between the sexes is that males lead noisier lives than females and consequently have

TABLE I. Values of F and p for 2-factor ANOVA conducted on the log transform of the number of SOAEs per ear, with the summary statistics to the right.

Factor/means comparison	Number of SOAEs F	p
Subject Type	18.30	0.0001
Ear Difference	7.61	0.006
Subject Type×Ear Difference	0.67	0.572
Heterosexual ♀ vs Heterosexual ♂	35.39	0.0001
Heterosexual ♀ vs Homosexual+Bisexual ♀	9.35	0.003
Heterosexual ♂ vs Homosexual+Bisexual ♂	0.25	0.614

Subject type	Statistics	Number of SOAEs Left ear	Right ear
Heterosexual ♀	Mean	4.77	5.47
	N	57	57
	Stand. Error	0.53	0.58
Nonheterosexual ♀	Mean	2.98	3.20
	N	60	60
	Stand. Error	0.40	0.42
Nonheterosexual ♂	Mean	1.26	1.74
	N	62	62
	Stand. Error	0.24	0.30
Heterosexual ♂	Mean	1.59	1.98
	N	56	56
	Stand. Error	0.31	0.41

First three entries at left are results from a 2-factor ANOVA, with repeated measures on the ear factor.
Final three entries are results of means comparisons.
For ANOVA, number of SOAEs was log transformed as $\ln(X+1)$.
All 235 subjects included.
SOAEs from 550 to 9000 Hz included.
The degrees of freedom for the ear factor and the means comparisons were 1 and 231.
Where appropriate, the p values given are the Greenhouse–Geisser values.

more exposure-induced hearing loss (e.g., Axelsson *et al.*, 1987). This reanalysis contradicts that interpretation.

This reanalysis also sheds light on the issue of whether heterosexual females exhibit stronger overall SOAE power than nonheterosexual females (Fig. 2, top panel) solely because the former have more SOAEs (Fig. 1). As the entries at the bottom right of Table II reveal, the significance levels for the analyses of the measure power per SOAE were generally higher here than for the corresponding analyses of the full range of SOAEs (the differences between groups were larger and the variability was about the same). Taken together, the analyses over the two frequency ranges suggest that the larger number of SOAEs in heterosexual females does contribute somewhat to their having greater overall SOAE power than the nonheterosexual females. However, were it the sole basis for the differences among subject groups, the measure power per SOAE would not necessarily exhibit the same pattern across groups as seen for number of SOAEs, yet it does, both in Fig. 2 (bottom panel) and in the reanalyzed data on the SOAEs from 550 to 3000 Hz (Table II,

TABLE II. Values of F and p for various statistical comparisons, shown for three different measures of SOAEs, and two different frequency ranges of SOAEs, using only the subjects having SOAEs in both ears.

Factor/means comparison	Number of SOAEs F	p	Overall SOAE power F	p	Power/SOAE F	p
SOAE range: 550–9000 Hz						
Subject Type	5.26	0.002	3.67	0.014	2.46	0.066
Ear Difference	8.22	0.005	7.34	0.008	3.95	0.049
Subject Type×Ear Difference	0.36	0.780	0.49	0.687	0.97	0.407
Heterosexual ♀ vs Heterosexual ♂	11.92	0.0008	9.26	0.003	6.58	0.012
Heterosexual ♀ vs Homosexual+Bisexual ♀	5.36	0.022	3.68	0.057	2.46	0.119
Heterosexual ♂ vs Homosexual+Bisexual ♂	0.04	0.840	0.29	0.594	0.43	0.515
SOAE range: 550–3000 Hz						
Subject Type	4.70	0.004	4.64	0.004	3.80	0.012
Ear Difference	11.31	0.001	10.20	0.002	5.66	0.019
Subject Type×Ear Difference	0.27	0.851	0.04	0.991	0.13	0.941
Heterosexual ♀ vs Heterosexual ♂	8.55	0.004	9.36	0.003	8.19	0.005
Heterosexual ♀ vs Homosexual+Bisexual ♀	8.15	0.005	7.13	0.009	5.26	0.024
Heterosexual ♂ vs Homosexual+Bisexual ♂	0.01	0.944	0.01	0.926	0.04	0.847

First three rows of top and bottom sections are results of a two-factor ANOVA, with repeated measures on the ear factor.
Final three rows of each section are results of means comparisons.
For analysis, number of SOAEs was log transformed as $\ln(X+1)$.
For all variables, only subjects having SOAEs in both ears were included: $N=132$ and $N=120$ in top and bottom sections, respectively.
For the top section, the degrees of freedom for the ear factor and the means comparisons were 1 and 128; for the bottom section, they were 1 and 116.
Outcome for number of SOAEs when all 235 subjects are included is in Table I.
Where appropriate, the p values given are the Greenhouse–Geisser values.

TABLE III. The number (and *proportions*) of subjects having SOAEs, and the distribution of SOAEs across ears.

SOAEs in:	Heterosexual females	Homosexual females	Bisexual females	Bisexual males	Homosexual males	Heterosexual males	Total
Both ears	46	24	16	3	17	26	132
	0.807	*0.667*	*0.667*	*0.273*	*0.333*	*0.464*	
Left ear only	2	4	3	0	8	3	20
	0.035	*0.111*	*0.125*	*0.000*	*0.157*	*0.054*	
Right ear only	4	2	2	5	8	4	25
	0.070	*0.056*	*0.083*	*0.455*	*0.157*	*0.071*	
Neither ear	5	6	3	3	18	23	58
	0.088	*0.167*	*0.125*	*0.273*	*0.353*	*0.411*	
Total subjects	57	36	24	11	51	56	235
	1.000	*1.000*	*1.000*	*1.000*	*1.000*	*1.000*	

bottom). Accordingly, it seems appropriate to conclude that there is an inherent difference in the strength of the individual cochlear amplifiers across subject groups, a conclusion that is reinforced by the parallel differences in CEOAE strength (McFadden and Pasanen, 1998).

D. Proportions of subjects exhibiting SOAEs

The SOAE data are described in one final way in Table III, which shows the numbers (and proportions in italics) of subjects in each group having, and not having, SOAEs in the two ears. As can be seen, the proportion of subjects having SOAEs in both ears was greatest for the heterosexual females and markedly smaller for the heterosexual males, with the homosexual and bisexual females again being intermediate to the heterosexual females and heterosexual males. A three-factor loglinear analysis was performed on these data. The factors were sex (two levels), sexual orientation (three levels: heterosexual, homosexual, or bisexual), and SOAE expression (three levels: SOAEs in both ears, one ear only, or neither ear). Because the factors of sex and sexual orientation were considered fixed by the experimental design, the interaction between sex and sexual orientation was included in all models tested; that is, a logit model was employed. The overall effect of sex was highly significant, but sexual orientation did not alter SOAE expression significantly, nor did it alter the relationship between male and female subjects. That is, deleting the three-way interaction from the saturated model did not significantly alter the fit of the model, Chi-square $(4)=2.1$, $p=0.72$, nor did deleting sexual orientation, Chi-square $(8)=11.9$, $p=0.15$, but deleting the factor of sex from the model did significantly worsen the fit, Chi-square $(6)=31.3$, $p=0.00002$. Thus, males and females differed in their expression of SOAEs, but subjects with different sexual orientation did not. Even though the difference in SOAE prevalence between heterosexuals and homosexuals was not significant, we note that the pattern of the data has the nonheterosexual females placed between the heterosexual females and heterosexual males, just as was true with all the other OAE measures considered. [This analysis was done using Statistica (StatSoft, Inc., Tulsa, OK, 1994).]

E. Correlations between OAE measures

In an attempt to determine the degree of relationship among the various OAE measures obtained from these subjects, correlations were calculated for the 132 subjects having SOAEs in both ears (for the 550–9000 Hz data). SOAE number and overall SOAE power were correlated 0.87, overall SOAE power and power per SOAE were correlated 0.97, and CEOAE strength was correlated 0.76 with SOAE number, 0.79 with overall SOAE power, and 0.73 with power per SOAE. Two-ear averages were used for all of these OAE measures.[2] While SOAEs and CEOAEs are both regarded to be by-products of the hypothesized cochlear amplifiers (Davis, 1983), the underlying mechanisms are apparently not identical for the two. For example, dissociations are known to occur (e.g., Long and Tubis, 1988; McFadden and Pasanen, 1994). Those facts, plus the less-than-perfect correlations between CEOAEs and SOAEs, reveal that the SOAE data are not perfectly redundant to the CEOAE data, and there was value in our analyzing both sets of data. Observing the same pattern of results in the various SOAE measures as was obtained in the CEOAE data provides valuable assurance that the CEOAE findings were not a chance occurrence.

F. Effect sizes

Table IV contains estimates of the effect sizes observed here for selected pairings of subject categories (for the 550–9000 Hz data). Effect sizes were calculated by dividing the difference between the means of the two groups or conditions of interest by the square root of the weighted mean of their two variances. For the two measures of SOAE power, the calculations are based upon the 132 subjects having SOAEs in both ears; for the measure number of SOAEs, all 235 subjects were used. Also shown for comparison are the corresponding effect sizes extracted from the CEOAE data (McFadden and Pasanen, 1998) from all 237 of these same subjects (some of the values reported in Table I of the CEOAE paper contained errors in the second decimal place). For situations like these, Cohen (1992) suggested that effect sizes of 0.2, 0.5, and 0.8 could be viewed as small, medium, and large, respectively. Generally, the largest effect sizes

TABLE IV. Effect sizes for various comparisons.

Comparison	Number of SOAEs	Overall power of SOAEs	Power/SOAE	Strength of CEOAEs
Heterosexual ♀ vs Heterosexual ♂	0.98	0.72	0.60	0.76
Heterosexual ♀ vs Homosexual ♀	0.61	0.43	0.32	0.37
Heterosexual ♀ vs Bisexual ♀	0.45	0.32	0.31	0.44
Heterosexual ♀ vs Homosexual+Bisexual ♀	0.57	0.41	0.42	0.41
Homosexual ♀ vs Bisexual ♀	−0.18	−0.13	−0.02	0.09
Homosexual ♀ vs Homosexual ♂	0.52	0.08	0.07	0.54
Heterosexual ♂ vs Homosexual ♂	0.08	−0.23	−0.24	0.07
Heterosexual ♂ vs Bisexual ♂	0.28	0.12	−0.06	0.16
Heterosexual ♂ vs Homosexual+Bisexual ♂	0.12	−0.17	−0.21	0.03
Homosexual ♂ vs Bisexual ♂	0.25	0.32	0.15	0.26
Ear Differences for:				
Heterosexual ♀	0.17	0.20	0.22	0.26
Homosexual ♀	0.08	0.35	0.28	0.26
Bisexual ♀	0.05	0.25	0.33	0.04
Bisexual ♂	0.76	−0.25	−0.30	0.69
Homosexual ♂	0.15	0.26	0.17	0.31
Heterosexual ♂	0.14	0.17	0.12	0.20

Two-ear averages used for calculating top half of table.
N's for the four OAE measures were 235, 132, 132, and 237 from left to right.
For CEOAEs, averages across four click levels used throughout.

were obtained for the measure number of SOAEs, and the smallest for the measure power per SOAE. (Using the log-transformed values of the measure number of SOAEs produced effect sizes only slightly different from those shown.) All comparisons involving the bisexual males must be viewed with considerable caution given the small number of those subjects.

G. Use of oral contraceptives

Bancroft *et al.* (1991) reported that the levels of free testosterone in the plasma were substantially lower in women taking oral contraceptives than in normal-cycling women. Because far more of our heterosexual subjects (21/57) were using oral contraceptives than our homosexual (2/36) and bisexual (6/24) subjects, this fact constituted a possible explanation for our basic pattern of results. The higher androgen levels in the nonheterosexual females might have acted to suppress the cochlear amplifiers, thereby diminishing OAEs, meaning that the group differences observed in OAEs might be attributable largely or solely to the group difference in oral-contraceptive use. As a test of this possibility, SOAEs and CEOAEs were compared for users and nonusers of oral contraceptives both within and across subject categories. For all the measures of SOAEs, and for CEOAE magnitude, the results ran counter to the possible explanation. The OAE measures were uniformly *smaller* in the subjects using oral contraceptives than in the nonusers, although the differences were not statistically significant. It appears that this difference in lifestyle is not responsible for the differences in OAEs reported here for heterosexual and nonheterosexual females. Other auditory measures also show weak masculinizing effects among the users of oral contraception (see McFadden, 1999). In passing, we note that, for the users of oral contraceptives, the two measures of SOAE power were both slightly larger in the left ear instead of the standard pattern of being substantially larger in the right ear. SOAE number and CEOAE power did exhibit the traditional pattern of being larger in the right ear. This result may be attributable to the drugs acting differentially to reduce the strength of strong SOAEs more than weak SOAEs.

H. Choice of SOAE-detection criterion

As noted above, the values of the various parameters used in our objective SOAE-detection algorithm were selected to yield an acceptably conservative compromise between the two goals of detecting true SOAEs that are weak and not identifying as SOAEs events that are not SOAEs. The first goal was important because it was logically possible that the different subject groups would not have different numbers of SOAEs, but the SOAEs would just be weaker in some groups than others. Not having a sensitive procedure then would have opened us to the criticism that we had reached incorrect conclusions simply because of undercounting weak SOAEs. However, having a sensitive procedure clearly has consequences for the other SOAE measures used here. For example, a sensitive procedure might detect a relatively large number of weak SOAEs in ears also having large numbers of strong SOAEs (e.g., the ears of heterosexual females). That should affect very little the overall SOAE power calculated for each of those ears, but it would greatly diminish the power per SOAE that was calculated. To be concrete, it may be that the compression in the range seen in the bottom panel of Fig. 2 relative to the top panel is due largely to the decision to use parameters in the SOAE-detection algorithm that permitted the detection of weak SOAEs. As a test of this possibility, the detection criterion in the algorithm (see Sec. I) was increased from 5 s.d.s to 8 s.d.s, and the three measures (number of SOAEs, overall SOAE power, and power per SOAE) were again extracted and tested statistically. As a consequence of this parameter

change, the total number of identified SOAEs diminished from 584 to 482 for the heterosexual females and from 371 to 300 for the nonheterosexual females. Also, the number of subjects having SOAEs in both ears dropped from 132 to 121. Nevertheless, the general pattern of the data was still exactly the same as before, with the nonheterosexual females being intermediate to the heterosexual females and heterosexual males on all three OAE measures, and with no significant differences among the males. Also, no statistical comparison that was significant ceased to be. Of primary interest here is the fact that the differences between the heterosexual and nonheterosexual females for the two measures of SOAE power were indeed larger than with the more lax detection criterion. The means comparison between the heterosexual and nonheterosexual females was significant both for overall SOAE power, $F(1,117)=6.44$, $p=0.01$, and for power per SOAE, $F(1,117)=4.64$, $p=0.03$. That is, omitting the weakest of the SOAEs did in fact increase the differences in SOAE power between the heterosexual and nonheterosexual females. This exercise is not an argument for using a stricter criterion for detecting SOAEs because on the whole the 5-s.d. detection rule agreed better with experienced visual judgment in identifying SOAEs than did the 8-s.d. rule. Rather, the point is that reducing the contribution of weak SOAEs further strengthens the conclusion that the cochlear amplifiers are in fact stronger in heterosexual females than in nonheterosexual females.

III. DISCUSSION

To summarize, measures of both SOAE number and SOAE strength showed the same general pattern of results across subject groups as did the data on CEOAEs (McFadden and Pasanen, 1998) obtained from the same subjects. Namely, the SOAEs of homosexual and bisexual females were both less numerous and weaker than those of heterosexual females. On all SOAE measures, the homosexual and bisexual females were intermediate to heterosexual females and heterosexual males, just as for the CEOAEs, and again there were no statistically significant differences among the males. The nearly identical pattern of results for all these OAE variables strengthens considerably the argument that the cochleas of nonheterosexual females are different from those of heterosexual females, and are shifted in the male direction.

As was true for the CEOAE data (McFadden and Pasanen, 1998), there are multiple possible explanations for the fewer and weaker SOAEs in nonheterosexual females. However, any explanation that appeals to differences in the sizes and/or acoustical properties of the middle and outer ear systems of heterosexual and nonheterosexual females is inconsistent with the fact that CEOAE magnitude has been shown to be not dependent upon either ear-canal volume or compliance of the tympanic membrane in a (presumably largely) heterosexual population (Harkrider and McFadden, submitted). Also, any explanation that suggests there is something different about the power spectra extracted from subjects of different orientations that makes our algorithm for detecting SOAEs differentially effective is inconsistent with the fact that the noise floors of the smoothed spectra extracted from the different subject groups were essentially identical (see Sec. I).

While OAEs do appear to be reasonably stable through life, they can be diminished by a number of agents including exposure to intense noise (Norton *et al.*, 1989) and various drugs (e.g., McFadden and Plattsmier, 1984; Long and Tubis, 1988; McFadden and Pasanen, 1994). So perhaps there is something about the lifestyles of homosexual and bisexual females that leads to weaker OAEs than those exhibited by heterosexual females. We wish to emphasize that precautions were taken to ensure that none of the agents currently known to diminish OAEs were operating in our subjects. To be included in the experiment, subjects had to pass a hearing screening test, and questions were asked about recent noise exposure and drug use prior to any OAE testing. Further, a reanalysis of the SOAE data using only the SOAEs between 550 and 3000 Hz produced the same statistical conclusions as the analysis of the full set of data. Finally, the possibility that the differences observed across subject groups were attributable to differential use of oral contraceptives by the heterosexual and nonheterosexual females was tested and rejected. Of course, there may yet be unrecognized lifestyle agents or activities that are responsible for the diminished OAEs, and, if so, one might expect these relationships between OAEs and sexual orientation not to hold across cultures and eras. Because of the difficulties inherent to ruling out all possible differences in lifestyle, explanations of this sort are likely to continue to be a logical possibility for some time.

We believe that the correct explanation will ultimately prove to be one that involves hormonal mechanisms of development and sexual differentiation, possibly operating during prenatal life. The evidence is largely circumstantial, but compelling, in part because of the way it agrees with facts and evidence from other domains. The logic of the argument is as follows:

(1) For mammals, it is commonly believed that the default phenotype is female, and the production of a male fetus requires exposure to high levels of androgens during the second trimester of prenatal development (cf. Fitch *et al.*, 1998).
(2) Among the many masculinizing effects of androgens on the body, brain, and behavior of a male fetus is a change in the choice of sexual partner from the default choice of male to that of female. The fact that homosexual females also prefer females as sexual partners is in accord with the idea that some brain site(s) responsible for sexual preference have been masculinized at some point in development, perhaps prenatally.
(3) Females typically have more SOAEs and stronger CEOAEs than males, suggesting that the process of masculinization (namely, exposure to high levels of androgens) reduces the strength of the cochlear amplifiers in males, and thereby diminishes their OAEs. Accordingly, the existence of OAEs that are displaced in the male direction in homosexual and bisexual females suggests that the same processes that masculinized whatever brain

centers are responsible for sexual preference also partially masculinized the cochleas of the nonheterosexual females.

(4) Females having male co-twins (opposite-sex dizygotic, or OSDZ, females) have OAEs that are more like those of males than those of other females (McFadden, 1993; McFadden *et al.*, 1996). A plausible explanation for this difference is that the OSDZ females were exposed to higher-than-normal levels (for females) of androgens prenatally as a consequence of their having shared their intrauterine environment with a male co-twin. Parallel masculinizing effects of this sort are well-known in other mammals (vom Saal, 1989; Clark and Galef, 1998).

(5) The same patterns of sex and ear differences that exist in adults also exist in newborns and children (Burns *et al.*, 1992; Norton, 1992) even though the levels of sex hormones are the same in the two sexes at birth and then again after about two months of age (Smail *et al.*, 1981).

(6) The available longitudinal data suggest that OAEs remain reasonably constant through life (Kemp *et al.*, 1986; Franklin *et al.*, 1992; Burns *et al.*, 1993, 1994; Prieve *et al.*, 1993).

(7) From (5) and (6), it is perfectly reasonable to believe that the differences that exist in the OAEs of both adult OSDZ females and adult nonheterosexual females also existed earlier in life. If those differences existed at birth, then they must be the product of masculinizing mechanisms that operated prenatally. There is no compelling evidence that androgen levels are systematically higher in adult OSDZ females than in same-sex dizygotic females (Julie Aikten Harris, personal communication, 9/21/98) or higher in adult nonheterosexual females than in adult heterosexual females (e.g., Downey *et al.*, 1987; Dancey, 1990), implying that the diminished OAEs are a residual effect of an earlier exposure. Further, exposure to abnormally high levels of androgens prenatally is reported to increase the prevalence of homosexual behavior and fantasy in females having congenital adrenal hyperplasia (Dittman *et al.*, 1990a, b, 1992). In accord with this idea of a prenatal factor contributing strongly to homosexuality is the fact that many future homosexuals behave differently from future heterosexuals from a young age (Bailey and Zucker, 1995; Manosevitz, 1970).

Overall, it seems parsimonious to assume that some structures in the brains and cochleas of homosexual and bisexual females were masculinized prenatally as a consequence of overexposure to androgenic agents. At the very least, this hypothesis appears worthy of serious experimental consideration.

The explanation presented here emphasizes differences in exposure to androgens and ignores possible differences in exposure to estrogens. Ultimately, this will almost certainly prove to be too simplistic given the complex way these hormones function and interact through development. For the time being, however, it seems parsimonious to assume that the gain of the cochlear amplifiers is set to the default level in females, and some factor operates to reduce that setting in males, rather than the reverse. What is known about the effects of estrogen on OAEs is that SOAEs change only minimally through the course of both the menstrual cycle (Bell, 1992; Haggerty *et al.*, 1993; Penner, 1995) and pregnancy (Burns *et al.*, 1993).

When contemplating possible mechanisms for the hormonal effects being proposed, it may be tempting to imagine a developing female embryo, or its mother, producing too much androgen (or converting too much testosterone to estradiol) at some state in development, which masculinizes the fetus' cochleas, whatever brain centers are ultimately responsible for sexual preference, and presumably other parts of the brain and body. As an alternative to higher androgen concentrations *per se*, it is possible that certain brain sites in some developing female embryos are, for some reason, hypersensitive to androgen at some stage(s) in early development, and take up too much of it. At an intuitive level, this possibility appears to carry more potential for the aftereffects to be relatively localized structurally and functionally than does the idea of generally higher androgen concentrations. The great extent to which heterosexual and homosexual females are generally *similar*, physiologically and behaviorally, seems to imply relatively localized structural differences between them.

It is important to emphasize again that knowledge of an individual person's SOAEs (or CEOAEs) provides very little information about the person's sexual orientation. For all of the dependent variables reported here, the distributions of scores for homosexual and bisexual females overlap considerably those for heterosexual females. The differences reported here are group differences.

If it proves to be the case that the same processes that masculinize the brain also masculinize the cochlea, that will establish the ear as a potentially valuable, noninvasive window onto certain aspects of brain development and sexual differentiation.

ACKNOWLEDGMENTS

We greatly appreciate the help of the following with recruiting and scheduling subjects, and data collection, entry, and reduction: N. L. Callaway, S. Oropeza, L. Ramirez, S. Bratcher, C. Furche, A. Phan, M. Dittrich, L. Strother. Especially helpful during the planning stages of the experiment were S. E. Finn and J. M. Bailey. P. K. Randall helped with the loglinear analysis. J. C. Loehlin, E. M. Hull, and three anonymous reviewers provided numerous helpful comments on a preliminary version of this paper. This work was supported by a research grant from the National Institute on Deafness and other Communication Disorders (NIDCD 00153).

[1]Technically, level is the correct term for describing values expressed in decibels sound-pressure level. The term power is used here intentionally as a reminder that the SOAE detection algorithm operated on the power spectrum, and the calculations of SOAE strength were done entirely in power units, with only the final results being converted into decibels.

[2]Previously published correlations of this sort were generally smaller (McFadden *et al.*, 1996), presumably because of a restriction of the range caused by the less sophisticated nature of our earlier method of detecting SOAEs. The average number of SOAEs shown in Fig. 1 is about twice the number reported in our twins experiment (McFadden and Loehlin, 1995), and this is largely attributable to the quieter spectra and off-line analysis

made possible by the collection of 2 min of data from each ear. The use of detection criteria different from the 5.0 s.d.s used here did not alter the ordering of the groups.

Axelsson, A., Aniansson, G., and Costa, O. (**1987**). ''Hearing loss in school children,'' Scand. Audiol. **16**, 137–143.

Bailey, J. M., and Zucker, K. J. (**1995**). ''Childhood sex-typed behavior and sexual orientation: A conceptual analysis and quantitative review,'' Dev. Psych. **31**, 43–55.

Bancroft, J., Sherwin, B. B., Alexander, G. M., Davidson, D. W., and Walker, A. (**1991**). ''Oral contraceptives, androgens, and the sexuality of young women: II. The role of androgens,'' Arch. Sex. Behav. **20**, 121–135.

Bell, A. (**1992**). ''Circadian and menstrual rhythms in frequency variations of spontaneous otoacoustic emissions from human ears,'' Hearing Res. **58**, 91–100.

Burns, E. M., Arehart, K. H., and Campbell, S. L. (**1992**). ''Prevalence of spontaneous otoacoustic emissions in neonates,'' J. Acoust. Soc. Am. **91**, 1571–1575.

Burns, E. M., Campbell, S. L., and Arehart, K. H. (**1994**). ''Longitudinal measurements of spontaneous otoacoustic emissions in infants,'' J. Acoust. Soc. Am. **95**, 385–394.

Burns, E. M., Campbell, S. L., Arehart, K. H., and Keefe, D. H. (**1993**). ''Long-term stability of spontaneous otoacoustic emissions,'' Abst. Assoc. Res. Otolaryngol. **16**, 98(A).

Clark, M. M., and Galef, Jr., B. G. (**1998**). ''Effects of intrauterine position on the behavior and genital morphology of litter-bearing rodents,'' Develop. Neuropsych. **14**, 197–211.

Cohen, J. (**1992**). ''A power primer,'' Psychol. Bull. **112**, 155–159.

Dancey, C. P. (**1990**). ''Sexual orientation in women: An investigation of hormonal and personality variables,'' Biol. Psychol. **30**, 251–264.

Davis, H. (**1983**). ''An active process in cochlear mechanics,'' Hearing Res. **9**, 79–90.

Dittman, R. W., Kappes, M. H., Kappes, M. E., Börger, D., Stegner, H., Willig, R. H., and Wallis, H. (**1990a**). ''Congenital adrenal hyperplasia. I. Gender-related behavior and attitudes in female patients and sisters,'' Psychoneuroendocrinology **15**, 401–420.

Dittman, R. W., Kappes, M. H., Kappes, M. E., Börger, D., Meyer-Bahlburg, H. F. L., Stegner, H., Willig, R. H., and Wallis, H. (**1990b**). ''Congenital adrenal hyperplasia. II. Gender-related behavior and attitudes in female salt-wasting and simple-virilizing patients,'' Psychoneuroendocrinology **15**, 421–434.

Dittman, R. W., Kappes, M. E., and Kappes, M. H. (**1992**). ''Sexual behavior in adolescent and adult females with congenital adrenal hyperplasia,'' Psychoneuroendocrinology **17**, 153–170.

Downey, J., Ehrhardt, A. A., Schiffman, M., Dyrenfurth, I., and Becker, J. (**1987**). ''Sex hormones in lesbian and heterosexual women,'' Horm. Behav. **21**, 347–357.

Fitch, R. H., Cowell, P. E., and Denenberg, V. H. (**1998**). ''The female prototype: Nature's default?,'' Develop. Neuropsych. **14**, 213–231.

Franklin, D. J., McCoy, M. J., Martin, G. K., and Lonsbury-Martin, B. L. (**1992**). ''Test/retest reliability of distortion-product and transiently evoked otoacoustic emissions,'' Ear Hear. **13**, 417–429.

Haggerty, H. S., Lusted, H. S., and Morton, S. C. (**1993**). ''Statistical quantification of 24-hour and monthly variabilities of spontaneous otoacoustic emission frequency in humans,'' Hearing Res. **70**, 31–49.

Harkrider, A. W., and McFadden, D. (**1999**). ''Role of the outer and middle ears in the large variability in normative data on click-evoked otoacoustic emissions,'' Ear Hear. (submitted).

Harris, J. A. (**1998**). Personal communication.

Kemp, D. T. (**1978**). ''Stimulated acoustic emissions from within the human auditory system,'' J. Acoust. Soc. Am. **64**, 1386–1391.

Kemp, D. T. (**1979**). ''Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea,'' Arch. Oto-Rhino-Laryngol. **224**, 37–45.

Kemp, D. T., Bray, P., Alexander, L., and Brown, A. M. (**1986**). ''Acoustic emission cochleography—Practical aspects,'' Scand. Audiol. Suppl. **25**, 71–95.

Lindman, H. R. (**1974**). *Analysis of Variance in Complex Experimental Designs* (Holden-Day, San Francisco).

Long, G. R., and Tubis, A. (**1988**). ''Modification of spontaneous and evoked otoacoustic emissions and associated psychoacoustic microstructure by aspirin consumption,'' J. Acoust. Soc. Am. **84**, 1343–1353.

Manosevitz, M. (**1970**). ''Early sexual behavior in adult homosexual and heterosexual males,'' J. Abnormal Psych. **76**, 396–402.

McFadden, D. (**1993**). ''A masculinizing effect on the auditory systems of human females having male co-twins,'' Proc. Natl. Acad. Sci. USA **90**, 11900–11904.

McFadden, D. (**1998**). ''Sex differences in the auditory system,'' Develop. Neuropsych. **14**, 261–298.

McFadden, D. (**1999**). ''Masculinizing effects of oral contraceptives on the auditory system,'' J. Acoust. Soc. Am. (in preparation).

McFadden, D., and Loehlin, J. C. (**1995**). ''On the heritability of spontaneous otoacoustic emissions: A twins study,'' Hearing Res. **85**, 181–198.

McFadden, D., and Pasanen, E. G. (**1994**). ''Otoacoustic emissions and quinine sulfate,'' J. Acoust. Soc. Am. **95**, 3460–3474.

McFadden, D., and Pasanen, E. G. (**1998**). ''Comparison of the auditory systems of heterosexuals and homosexuals: Click-evoked otoacoustic emissions,'' Proc. Natl. Acad. Sci. USA **95**, 2709–2713.

McFadden, D., and Plattsmier, H. S. (**1984**). ''Aspirin abolishes spontaneous oto-acoustic emissions,'' J. Acoust. Soc. Am. **76**, 443–448.

McFadden, D., Loehlin, J. C., and Pasanen, E. G. (**1996**). ''Additional findings on heritability and prenatal masculinization of cochlear mechanisms: Click-evoked otoacoustic emissions,'' Hearing Res. **97**, 102–119.

Norton, S. J. (**1992**). ''The effects of being a newborn on otoacoustic emissions,'' J. Acoust. Soc. Am. Suppl. 1 **91**, 2409.

Norton, S. J., Mott, J. B., and Champlin, C. A. (**1989**). ''Behavior of spontaneous otoacoustic emissions following intense acoustic stimulation,'' Hearing Res. **38**, 243–258.

Pasanen, E. G., and McFadden, D. (**1996**). ''An automated off-line procedure for detecting SOAEs from a power spectrum,'' J. Acoust. Soc. Am. **100**, 2628(A).

Pattatucci, A. M., and Hamer, D. H. (**1995**). ''Development and familiality of sexual orientation in females,'' Behav. Gen. **25**, 407–420.

Penner, M. J. (**1995**). ''Frequency variation of spontaneous otoacoustic emissions during a naturally occurring menstrual cycle, amenorrhea, and oral contraception: A brief report,'' Ear Hear. **16**, 428–432.

Penner, M. J., Glotzbach, L., and Huang, T. (**1993**). ''Spontaneous otoacoustic emissions: Measurement and data,'' Hearing Res. **68**, 229–237.

Prieve, B. A., Gorga, M. P., Schmidt, A., Neely, S., Peters, J., Schulte, L., and Jesteadt, W. (**1993**). ''Analysis of transient-evoked otoacoustic emissions in normal-hearing and hearing-impaired ears,'' J. Acoust. Soc. Am. **93**, 3308–3319.

Probst, R., Lonsbury-Martin, B. L., and Martin, G. K. (**1991**). ''A review of otoacoustic emissions,'' J. Acoust. Soc. Am. **89**, 2027–2067.

Smail, P. J., Reyes, F. I., Winter, J. S. D., and Faiman, C. (**1981**). ''The fetal hormonal environment and its effect on the morphogenesis of the genital system,'' in *Pediatric Andrology*, edited by S. J. Kogan and E. S. E. Hafez (Martinus Nijhoff, The Hague), pp. 9–19.

Smurzynski, J., and Probst, R. (**1998**). ''The influence of disappearing and reappearing spontaneous otoacoustic emissions on one subject's threshold microstructure,'' Hearing Res. **115**, 197–205.

Talmadge, C. L., Long, G. R., Murphy, W. J., and Tubis, A. (**1993**). ''New off-line method for detecting spontaneous otoacoustic emissions in human subjects,'' Hearing Res. **71**, 170–182.

vom Saal, F. S. (**1989**). ''Sexual differentiation in litter-bearing mammals: Influence of sex of adjacent fetuses in utero,'' J. Animal Sci. **67**, 1824–1840.

Whitehead, M. L. (**1991**). ''Slow variations of the amplitude and frequency of spontaneous otoacoustic emissions,'' Hearing Res. **53**, 269–280.

Zwicker, E. (**1990**). ''On the frequency separation of simultaneously evoked otoacoustic emissions' consecutive extrema and its relation to cochlear traveling waves,'' J. Acoust. Soc. Am. **88**, 1639–1641.

Nonlinear active force generation by cochlear outer hair cell

Alexander A. Spector
Department of Biomedical Engineering and Center for Computational Medicine and Biology, The Johns Hopkins University, Baltimore, Maryland 21205

William E. Brownell
Bobby R. Alford Department of Otorhinolaryngology and Communicative Sciences, Baylor College of Medicine, Houston, Texas 77030

Aleksander S. Popel
Department of Biomedical Engineering and Center for Computational Medicine and Biology, The Johns Hopkins University, Baltimore, Maryland 21205

(Received 19 March 1998; revised 14 December 1998; accepted 24 December 1998)

We analyze the nonlinear behavior of the longitudinal and circumferential components of the active force generated by the outer hair cell wall in response to changes of its transmembrane potential. We treat the material of the wall as electroelastic, linear orthotropic in terms of strains and as nonlinear in terms of the transmembrane potential. To describe the nonlinear behavior of the active force versus the transmembrane potential, we use two (Boltzmann and simple exponential) types of approximation. We estimate free parameters of these approximations by combining the previously reported passive stiffnesses with the active strains measured in the microchamber experiment. We analyze the sensitivity of the estimated parameters corresponding to changes of the cell axial stiffness, a characteristic independently measured by several groups. We also study the effect of combining the active strains measured in the microchamber experiment with those measured in the whole cell recording experiment. We show agreement between our prediction of the active force and measurements in the whole cochlea and in isolated cells. © *1999 Acoustical Society of America.* [S0001-4966(99)01104-2]

PACS numbers: 43.64.Ld, 43.64.Kc, 43.64.Bt [RDF]

INTRODUCTION

Cochlear outer hair cells are believed to be the major active component in the mammalian hearing process. These cells can generate active forces that dramatically increase the amplitudes of the basilar membrane and reticular lamina vibration, resulting in the high amplification and frequency selectivity of the ear. The active force generation is intrinsically related to changes in cell length. These length changes, called electromotility, were first observed in isolated cells by Brownell *et al.* (1985). Further studies (Ashmore, 1987; Santos-Sacchi, 1988; Dallos *et al.*, 1993) showed that changes in cell length are primarily determined by changes in the transmembrane potential of the cell wall. Under physiological conditions, the transmembrane potential changes are caused by the opening of the mechano-electrical transduction channels located on the stereocilia bundle at the top of the cell. The channel's opening is related to the stereocilia bending. This bending is a result of the general cochlea mechanics where the basilar and tectorial membranes move with respect to each other.

The detailed observations of outer hair cell motility versus the transmembrane potential developed in isolated cells have revealed specific features of this phenomenon. The electromotile length changes are nonlinear and asymmetric with respect to the cell's resting potential under both *in vivo* and experimental conditions. There is a quasilinear range around the resting potential of the length dependence on the transmembrane potential. Outside this range the length changes are nonlinear with a monotonic decrease of the slope that can be interpreted as saturation in the regions of extreme depolarization and hyperpolarization. The dependence of the cell radius on the transmembrane potential has similar features (Hallworth *et al.*, 1993). It is necessary to emphasize that, although the *in vivo* range of the transmembrane potential is relatively narrow, the nonlinearity of the outer hair cell's length and radius changes is important for the understanding of the electromotility mechanism and its molecular basis.

Hallworth (1995, 1997) developed an experimental technique for measuring the active forces in an isolated outer hair cell. In this experiment, the cell is pressed against a fiber of high rigidity. Another version of this technique was developed by Iwasa and Adachi (1997). Xue *et al.* (1993) estimated the active force produced by a single outer hair cell by using *in vivo* data on the basilar membrane stiffness (Olson and Mountain, 1991) and the membrane displacements under the outer hair cell action.

Two approaches have been developed for the mathematical modeling of electromotility and the active force generation. The molecular motor approach explicitly includes parameters and relationships for an elementary active motor which allows the derivation of the active strains and forces in terms of the motor's characteristics (Dallos *et al.*, 1993; Iwasa, 1994; Iwasa and Adachi, 1997). A phenomenological approach based on linear piezoelectric relationships has been proposed by Steele *et al.* (1993) and Tolomeo and Steele (1995). In their model, the outer hair cell active be-

havior was characterized by effective coefficients estimated from the electromotility experiments.

In the present paper, we develop a nonlinear version of the phenomenological approach that is based on electroelastic relationships. The relationships we use reduce to those from Steele *et al.* (1993) and Tolomeo and Steele (1995) if we linearize the nonlinear terms at the point of the resting potential. We prefer the term ''electroelastic'' to ''piezoelectric'' because it more precisely characterizes the electromechanical coupling that occurs in the outer hair cell.

In our model, we use voltage-dependent approximations of the nonlinear terms responsible for the active forces. The free parameters of the approximations of each of two components of the active force are derived by using the values of the slopes at the point of the resting potential and the limiting values corresponding to the saturation regions. These values are estimated by using the previously reported passive stiffnesses in combination with the active strains measured in the micropipet experiment. First, we determine the free parameters for a Boltzmann-type approximation that is smooth up to the first derivative and skew-symmetric with respect to a certain point. Boltzmann-type approximations have been used for the treatment of the experimental data on the electromotile displacements.

In our approach, the Boltzmann-type approximation does not have any statistical foundation, but it is used as a function with certain approximating properties. For this reason, we also use an alternative simple exponential approximation that can capture the major features of the electromotility phenomenon. We estimate the free parameters of this approximation in a manner similar to what we do for the Boltzmann-type function.

For both approximations, we find that the nonlinearity is higher in the hyperpolarization region than in the depolarization region. Also, the limiting value [both for the longitudinal (along the cylindrical cell) and circumferential (around the cell) component of the active force] in the region of the extreme depolarization is greater than the corresponding limiting value in the region of the extreme hyperpolarization.

Estimation of the parameters of the active force is based on the values of the passive stiffnesses calculated by using experimental data on the cell axial stiffness. We use average data from several independent groups (Holley and Ashmore, 1988; Russell and Schauz, 1995; Ulfendahl *et al.*, 1998) falling in a relatively narrow range. To reflect measurements quantitatively different from those mentioned (Hallworth, 1997; Iwasa and Adachi, 1997), we present a comprehensive analysis of the active force parameters variation depending on the axial stiffness of the outer hair cell across a broad range of values.

In addition to the parameter estimation based on the active strains from the microchamber experiment only, we consider two possible alternative ways of treating the electromotility experimental data. In the first, we combine information on the measurements of the circumferential component of the active strain in the microchamber experiment with information on the measurements of the longitudinal component in the whole cell recording experiment. In the second, we combine information on the measurements of the longitudinal component in each of the two experiments. According to these alternative approaches, we re-estimate the parameters of the active force and demonstrate their small changes with respect to the data from the microchamber experiment only.

We show agreement of our estimates of the active force with experimental data obtained both in the whole cochlea (Xue *et al.*, 1993) and in isolated cells (Hallworth, 1997; Iwasa and Adachi, 1997).

I. DETERMINATION OF THE NONLINEAR VOLTAGE-DEPENDENT ACTIVE FORCES

A. General relationships

To determine the active forces generated by a single outer hair cell, we assume the material of the wall to be electroelastic, linear orthotropic in terms of strains and nonlinear in terms of the transmembrane potential change. In such a case, the longitudinal and circumferential components of the active force $f_x(\Delta V)$ and $f_\theta(\Delta V)$ are determined by the following equations:

$$N_x=(C_{11}\epsilon_x+C_{12}\epsilon_\theta)+f_x(\Delta V), \tag{1}$$

$$N_\theta=(C_{12}\epsilon_x+C_{22}\epsilon_\theta)+f_\theta(\Delta V). \tag{2}$$

Here N_x and N_θ are the components of the in-plane resultant forces, C_{11}, C_{12}, and C_{22} are the generalized stiffnesses of the outer hair cell wall, ϵ_x and ϵ_θ are the components of the strain, and ΔV is the transmembrane potential change with respect to the resting potential. In Eqs. (1) and (2), the forces f_x and f_θ have the meaning of active forces produced by the outer hair cell wall under isometric conditions. It follows from Eqs. (1) and (2) if we set $\epsilon_x=\epsilon_\theta=0$. Expressions of this type have been used for several models of the outer hair cell wall (Iwasa and Chadwick, 1992; Dallos *et al.*, 1993; Tolomeo and Steele, 1995; Iwasa and Adachi, 1997).

We analyze the nonlinear voltage-dependent behavior of the active forces f_x and f_θ by using the electromotile strains ϵ_x and ϵ_θ measured in the microchamber experiment (Dallos *et al.*, 1993). We also use the stiffnesses C_{11}, C_{12}, and C_{22} previously estimated by Spector *et al.* (1998b). The electromotile curves $\epsilon_x(\Delta V)$ and $\epsilon_\theta(\Delta V)$ exhibit characteristic features: a quasilinear range around the resting potential and a nonlinear behavior outside this range accompanied by saturation for the extreme depolarization and hyperpolarization regions (Dallos *et al.*, 1993; Santos-Sacchi, 1988). The limiting values of both components of the electromotile displacement corresponding to the extreme depolarization region are considerably different from those in the extreme hyperpolarization region.

Although the total volume inside the cell is preserved under the conditions of the microchamber experiment, the partial volumes of the included and excluded parts of the cell are free to change. It means that, opposite to the conditions of the whole cell recording experiment, the components e_x and e_θ for one part (included or excluded) of the cell are not linearly related. In this regard, they can be considered independent. Obviously, two pairs of the strain components that correspond to two (included and excluded) parts of the cell are not independent, and the corresponding constraint expresses the total volume preservation.

We use the experimental data for the case when two parts of the cell have equal lengths. Following Dallos *et al.* (1993, p. 313) and Tolomeo and Steele (1995), we neglect pressure changes caused by the electromotile strains. It means that we neglect the mechanical forces occurring under the conditions of the microchamber experiment, and therefore the components of the active force can be determined by the equations

$$-(C_{11}\epsilon_x + C_{12}\epsilon_\theta) = f_x(\Delta V), \tag{3}$$

$$-(C_{12}\epsilon_x + C_{22}\epsilon_\theta) = f_\theta(\Delta V). \tag{4}$$

The active force production per unit of the transmembrane potential change is an important parameter for the characterization of the outer hair cell/basilar membrane and outer hair cell/reticular lamina interactions. This parameter can be described by using the derivatives of the active forces components with respect to the transmembrane potential change. Because of the geometric symmetry of the case under consideration, there is no effect of the pressure change on these derivatives calculated at the point of the resting potential. The corresponding coefficients derived strictly from Eqs. (1) and (2) have the following form:

$$e_x = \frac{df_x}{d\Delta V}(\Delta V = 0) = -\left[C_{11}\frac{d\epsilon_x}{d\Delta V}(\Delta V = 0) + C_{12}\frac{d\epsilon_\theta}{d\Delta V}(\Delta V = 0)\right], \tag{5}$$

$$e_\theta = \frac{df_\theta}{d\Delta V}(\Delta V = 0) = -\left[C_{12}\frac{d\epsilon_x}{d\Delta V}(\Delta V = 0) + C_{22}\frac{d\epsilon_\theta}{d\Delta V}(\Delta V = 0)\right]. \tag{6}$$

The derivatives of the electromotile strains in Eqs. (5) and (6) are extracted from the electromotile curves. The cells studied in all experiments that we use for our analysis have the lengths falling within the range 50 μm–70 μm. Therefore the effect of the cell length variation should not be significant. An extrapolation of the obtained results to the short cells range needs an additional analysis, however, experimental data show no (Hallworth, 1997) or moderate (Iwasa and Adachi, 1997) effect of the cell length on the value of the active isometric force.

As follows from Eqs. (1) and (2), the components of the active force have the limits f_x^+ and f_θ^+ in the region of the extreme depolarization ($\Delta V \to \infty$) and f_x^- and f_θ^- in the region of extreme hyperpolarization ($\Delta V \to -\infty$). The limiting values of the active force are expressed in terms of the limiting values of the electromotile strains ϵ_x^+, ϵ_θ^+, ϵ_x^-, and ϵ_θ^- by the following equations:

$$f_x^+ = -(C_{11}\epsilon_x^+ + C_{12}\epsilon_\theta^+), \quad f_x^- = -(C_{11}\epsilon_x^- + C_{12}\epsilon_\theta^-), \tag{7}$$

$$f_\theta^+ = -(C_{12}\epsilon_x^+ + C_{22}\epsilon_\theta^+), \quad f_\theta^- = -(C_{12}\epsilon_x^- + C_{22}\epsilon_\theta^-). \tag{8}$$

B. Boltzmann-type approximation

The Boltzmann-type functions give approximations that are smooth up to the first derivative and skew-symmetric with respect to a certain point. This type of approximation has been commonly used for the analysis of outer hair cell electromotility (e.g., Dallos *et al.*, 1993; Gale and Ashmore, 1994). We use Boltzmann-type approximations for the components of the active force in the following form:

$$f_x(\Delta V) = a_x\frac{\exp[b_x(\Delta V - \Delta V_x^0)] - 1}{\exp[b_x(\Delta V - \Delta V_x^0)] + 1} - c_x, \tag{9}$$

$$f_\theta(\Delta V) = a_\theta\frac{\exp[b_\theta(\Delta V - \Delta V_\theta^0)] - 1}{\exp[b_\theta(\Delta V - \Delta V_\theta^0)] + 1} - c_\theta. \tag{10}$$

The parameters a_x, b_x, c_x, ΔV_x^0 and a_θ, b_θ, c_θ, ΔV_θ^0 are not independent. The first four parameters are related by the equation

$$f_x(\Delta V = \Delta V_x^0) = 0. \tag{11}$$

The second four parameters satisfy the equation

$$f_\theta(\Delta V = \Delta V_\theta^0) = 0. \tag{12}$$

Equations (11) and (12) are consequences of the skew-symmetry of functions $f_x(\Delta V)$ and $f_\theta(\Delta V)$ with respect to the points $\Delta V = \Delta V_x^0$ and $\Delta V = \Delta V_\theta^0$, respectively.

To find the free parameters of approximations (9) and (10), we use three pairs of conditions. We require functions (9) and (10) to have the same slopes (at the point of the resting potential) and the same limiting values that are prescribed by Eqs. (5)–(8). We derive the parameters of our approximations and express them in terms of f_x^+, f_θ^+, f_x^-, f_θ^-, e_x, and e_θ. Considering $\Delta V \to \infty$ and $\Delta V \to -\infty$, we obtain from Eqs. (9) and (10)

$$a_x = \frac{f_x^+ - f_x^-}{2}, \quad c_x = -\frac{f_x^+ + f_x^-}{2}, \tag{13}$$

$$a_\theta = \frac{f_\theta^+ - f_\theta^-}{2}, \quad c_\theta = -\frac{f_\theta^+ + f_\theta^-}{2}. \tag{14}$$

Taking into account Eqs. (13) and (14), Eqs. (11) and (12) result in the following coordinates of the skew-symmetry points:

$$\Delta V_x^0 = \ln\left(-\frac{f_x^+}{f_x^-}\right), \quad \Delta V_\theta^0 = \ln\left(-\frac{f_\theta^+}{f_\theta^-}\right). \tag{15}$$

Theoretically, the positions of the skew-symmetry point for the longitudinal and circumferential components are different ($\Delta V_x^0 \neq \Delta V_\theta^0$). However, the calculations below show that this difference is very small. Differentiating the right hand sides in Eqs. (9) and (10), we derive

$$e_x = -b_x\frac{f_x^+ f_x^-}{f_x^+ - f_x^-}, \quad e_\theta = -b_\theta\frac{f_\theta^+ f_\theta^-}{f_\theta^+ - f_\theta^-}. \tag{16}$$

Therefore, we can express the parameters b_x and b_θ by the equations

$$b_x = -e_x\frac{f_x^+ - f_x^-}{f_x^+ f_x^-}, \quad b_\theta = -e_\theta\frac{f_\theta^+ - f_\theta^-}{f_\theta^+ f_\theta^-}. \tag{17}$$

C. Simple exponential approximation

We consider another approximation of the active force which is consistent with the major features of the electromotile behavior of the outer hair cell and examine the difference in the predictions from two approximations. We use a simple exponential approximation in the following form:

$$f_x(\Delta V)=\begin{cases} a_x^+[1-\exp(-b_x^+\Delta V)] & \text{when } \Delta V\geqslant 0 \\ a_x^-[1-\exp(-b_x^-\Delta V)] & \text{when } \Delta V<0 \end{cases} \tag{18}$$

$$f_\theta(\Delta V)=\begin{cases} a_\theta^+[1-\exp(-b_\theta^+\Delta V)] & \text{when } \Delta V\geqslant 0 \\ a_\theta^-[1-\exp(-b_\theta^-\Delta V)] & \text{when } \Delta V<0 \end{cases}. \tag{19}$$

Obviously, we have

$$f_x^+=a_x^+,\ f_x^-=a_x^-,\ f_\theta^+=a_\theta^+,\ f_\theta^-=a_\theta^-. \tag{20}$$

Functions (9) and (10) are not automatically smooth at the point $\Delta V=0$ (the resting potential). We can provide smoothness of these functions by using the following equations relating the free parameters of our approximations:

$$b_x^-=\frac{a_x^+b_x^+}{a_x^-},\quad b_\theta^-=\frac{a_\theta^+b_\theta^+}{a_\theta^-}. \tag{21}$$

It can be easily checked that functions f_x and f_θ defined by Eqs. (18) and (19) under condition (20) have equal slopes to the right and to the left from the point $\Delta V=0$. The coefficients of the active force production are expressed by the formulas

$$e_x=a_x^+b_x^+=a_x^-b_x^-,\quad e_\theta=a_\theta^+b_\theta^+=a_\theta^-b_\theta^-. \tag{22}$$

II. ANALYSIS OF THE ACTIVE FORCE BEHAVIOR

Our estimates of the free parameters of the proposed approximations are based on the results of the microchamber experiment. We use graphs for the length and radius changes versus the command voltage varying in a broad range (Dallos *et al.*, 1993, Fig. 7). These graphs correspond to the symmetric case when the lengths of the parts of the cell excluded from the microchamber and included in it are equal. The graphs were obtained by Dallos *et al.* (1993) as a result of an extrapolation of actual measurements done for different ratios of those lengths. Each of the two graphs gives us the slope at the point of the resting potential (e_x or e_θ) and the limiting values corresponding to extremely positive (e_x^+ or e_θ^+) and extremely negative (e_x^- or e_θ^-) command voltages. We calculate the following parameters necessary for our estimates

$$\frac{d\epsilon_x}{d\Delta V}(\Delta V=0)\approx -0.22V^{-1},\quad \frac{d\epsilon_\theta}{d\Delta V}(\Delta V=0)\approx 0.09V^{-1}, \tag{23}$$

$$\epsilon_x^+\approx -0.048,\ \epsilon_x^-\approx 0.012,\ \epsilon_\theta^+\approx 0.012,\ \epsilon_\theta^-\approx -0.0026. \tag{24}$$

Substituting the data from Eqs. (23) and (24) into Eqs. (5)–(8), we obtain the starting parameters for both approximations. Equations (13)–(17) give us the free parameters of the Boltzmann-type approximation. By using Eqs. (20) and (21), we calculate the free parameters of the exponential approximation.

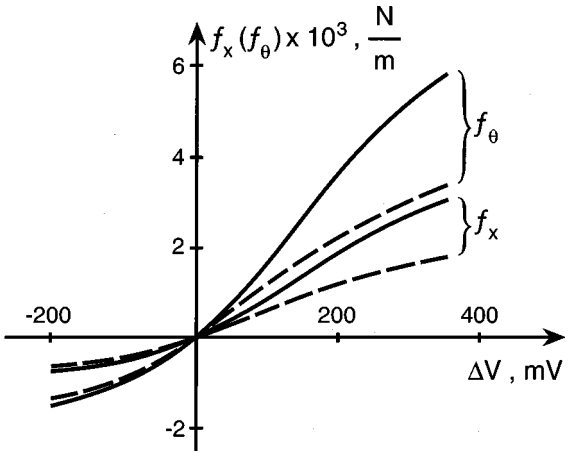


FIG. 1. Two components of the active force versus changes of the transmembrane potential. The solid lines correspond to the Boltzmann-type approximation, and the dashed lines correspond to the simple exponential approximation.

The longitudinal and circumferential components of the active force versus changes of the transmembrane potential are shown in Fig. 1. The solid and dashed lines correspond to the Boltzmann-type and the simple exponential approximations, respectively. The Boltzmann-type approximation gives greater absolute values of both components of the active force throughout the whole range of variation of the transmembrane potential. Within the hyperpolarization region ($\Delta V<0$) two approximations lead to close results. Within the depolarization region ($\Delta V>0$), these approximations are close in the quasilinear range and in the right saturation region. In the intermediate area, the difference between the two approximations reaches about 70%. The effect of the nonlinearity of the active force is much more substantial in the hyperpolarization area, and can be observed for the approximations of both types. The Boltzmann-type approximation starts saturating at lower absolute values of the transmembrane potential. This function does not change its curvature in the quasilinear range opposite to the simple exponential approximation.

The coefficient e_x of the active force production is of primary physiological importance because it corresponds to the force acting from the outer hair cell on the basilar membrane and the reticular lamina. We compared our estimates of these two coefficients with theoretical results obtained by using the in-plane stiffnesses derived from the isotropic model (Iwasa and Chadwick, 1992) and with experimental results (Xue *et al.*, 1993) that were based on direct measurements of the basilar membrane displacements in response to the outer hair cell action (Table I). As shown, our estimate of the longitudinal active force agrees well with the experimental data of Xue *et al.* (1993). Our estimate of the coefficient e_x is about two times smaller than that in the isotropic model.

We also compared our data with those from Steele's *et al.* (1993) and Tolomeo and Steele's (1995) orthotropic model where free parameters were estimated based on a different set of independent experiments. Our value of the co-

TABLE I. Coefficients characterizing the active force generation per unit of the transmembrane potential.

Coefficient of the active force	Present anisotropic model	Isotropic model[a]	Experimental result in the whole cochlea[b]
$e_x = df_x/d(\Delta V)$, N/Vm	6.8×10^{-3}	11×10^{-3}	6.4×10^{-3}
$e_\theta = df_\theta/d(\Delta V)$, N/Vm	13×10^{-3}	7×10^{-3}	

[a]Based on the isotropic membrane model of Iwasa and Chadwick (1992).
[b]Experimental results of Xue *et al.* (1993).

efficient e_x is about four times greater than the latest estimate of Tolomeo and Steele (1998). We reported previously (Spector *et al.*, 1998b) that our estimates of the in-plane Young's moduli and Poisson's ratios are close to those in Steele's *et al.* (1993) and Tolomeo and Steele's (1995) model. However, the generalized in-plane stiffnesses of the outer hair cell wall, C_{11}, C_{12}, and C_{22}, are very sensitive to the variation of Poisson's ratios. Small differences between Poisson's ratios can result in substantial differences between the corresponding stiffnesses C_{11}, C_{12}, and C_{22}. This is one reason for the difference between our data compared to Tolomeo and Steele's (1998); another reason is related to the interpretation of the experimental data on the outer hair cell axial stiffness. We use an average value based on data combined from several independent groups (Holley and Ashmore, 1988; Russell and Schauz, 1995; Ulfendahl *et al.*, 1998). In calculation of Steele *et al.* (1993) and Tolomeo and Steele (1995, 1998), a different pool of experimental data was used.

III. PARAMETER SENSITIVITY TO THE VARIATION OF THE CELL AXIAL STIFFNESS

The outer hair cell axial stiffness is an important experimental parameter for our estimation of the cell's mechanical properties. This parameter enters one of three equations that we used to estimate the wall in-plane stiffnesses (Spector *et al.*, 1998b). The effect of the cell axial stiffness on the active characteristics is determined by Eqs. (1) and (2). Several independent groups have measured the outer hair cell axial stiffness: while some of the results fall into a relatively narrow range (Holley and Ashmore, 1988; Russell and Schauz, 1995; Ulfendahl *et al.*, 1998), there are data giving the values 5–7 times (Hallworth, 1995, 1997) or more than an order of magnitude greater (Iwasa and Adachi, 1997). We previously reported an analysis of the variation of the in-plane stiffnesses corresponding to the changes of the axial stiffness in the interval 10^{-3} N/m–10^{-2} N/m. (Spector *et al.*, 1998b). Here, we use those data to analyze the effect of the cell axial stiffness on the active characteristics.

For both approximations, we determine the free parameters based on the values e_x, e_θ, f_x^+, f_x^-, f_θ^+, and f_θ^-, which we calculate by using the experimental information from electromotility experiments. In Figs. 2–4, we show the dependencies of these determining parameters on the cell axial stiffness γ_a. Figures 2, 3, and 4 represent, respectively, the coefficients e_x and e_θ; the limiting parameters f_x^+ and f_θ^+; and the limiting parameters f_x^- and f_θ^-. For our analysis here, we use an average value of the radial stiffness parameter that also enters our estimation of the in-plane stiffnesses. The radial stiffness parameter is the characteristic measured in the experiment with the micropipette aspiration of the outer hair cell wall (Sit *et al.*, 1997; Spector *et al.*, 1996, 1998a).

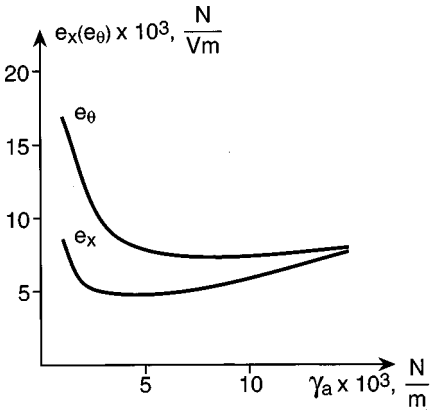


FIG. 2. The dependence of the coefficients of the active force generation on the axial stiffness of the outer hair cell γ_a.

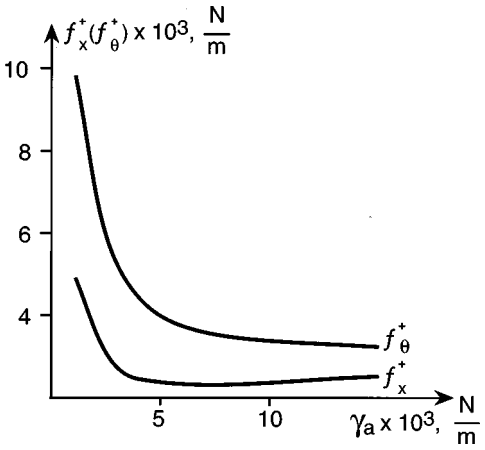


FIG. 3. The dependence of the extreme depolarization region values of the active force components on the axial stiffness of the outer hair cell γ_a.

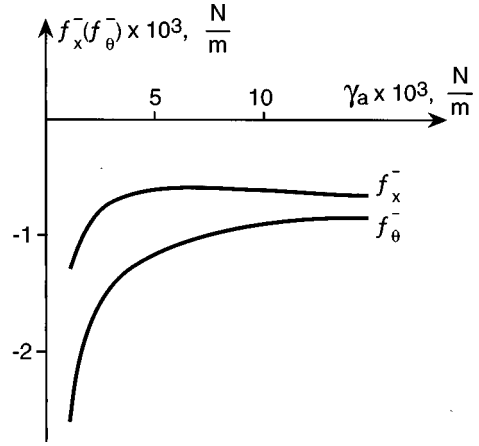


FIG. 4. The dependence of the extreme hyperpolarization region values of the active force components on the axial stiffness of the outer hair cell γ_a.

All six parameters in Figs. 2–4 exhibit similar behavior depending on the axial stiffness: decrease of the absolute value at low γ_a and slow variation at $\gamma_a \geqslant 6\times10^{-3}$ N/m. The total variation within the γ_a-interval under consideration is about two times for f_x^+ and f_θ^- and about three times for f_θ^+ and f_θ^-. The total variation of the coefficient e_θ is about 2.5 times. The variability of the most important characteristic, the coefficient e_x, is the weakest (about 90%) as compared to the other parameters. For all three pairs of the parameters (e_x, e_θ, f_x^+, f_θ^+; and f_x^-, f_θ^-,) the ratio of the circumferential component to the longitudinal one is very close to 2 when γ_a is between 10^{-3} N/m and 2×10^{-3} N/m. This ratio decreases when γ_a is about 10^{-2} N/m. Recently, the measurements of the active force generated by an isolated outer hair cell have been developed (Hallworth, 1995, 1997; Iwasa and Adachi, 1997). We can use our analysis of the active force variability with the axial stiffness for a better comparison of our predictions with the experimental data. Hallworth's (1997) latest experimental data give

$$\gamma_a \approx 6.6\times10^{-3}\ \text{N/m}, \quad e_x \approx 2\times10^{-3}\ \text{N/Vm}. \tag{25}$$

The value of the coefficient e_x corresponds to the isometric force that is practically independent of the cell length. This value is about 2.5 times lower than our estimate of the longitudinal component f_x of the active force (Fig. 3). Iwasa and Adachi's (1997) experimental results give

$$\gamma_a \approx 15\times10^{-3}\ \text{N/m}, \quad e_x \approx 4\times10^{-3}\ \text{N/Vm}. \tag{26}$$

These results show a moderate variation of the active force with the cell length (about 2–2.5 times within the broad interval 30 μm–90 μm). We chose the value of e_x that corresponds to the range 50 μm–70 μm. Estimating the c-stiffness and the e_x-coefficient for the same γ_a, we find that Iwasa and Adachi's (1997) result is about 1.8 times lower than our estimate.

IV. PARAMETER SENSITIVITY TO THE COMBINATION OF THE DATA FROM INDEPENDENT EXPERIMENTS WITH CELL ELECTRICAL STIMULATION

Above, we used data from the microchamber experiment to estimate the parameters of the active behavior of the outer hair cell. In that particular experiment, two components of the active strain can be considered independent. Under the conditions of the whole cell recording experiment (Santos-Sacchi, 1989), two components of the active strain are linearly related because of the cell volume preservation. We checked the consistency of the parameters estimated based on these two different techniques. To do this, we re-estimated the parameters by using a combination of the experimental data in which the longitudinal component of the active strain was taken from the whole cell recording experiment, and the circumferential component was taken from the microchamber experiment. After this, we compared the results with the original estimates that were based on the microchamber experiment only. As an additional way of testing the results, we analyzed the longitudinal component of the active force measured independently in the two experiments.

TABLE II. Differences between the active force generation coefficients obtained from different combination of experimental data. e_x, e_θ correspond to both components of the active force taken from the microchamber experiment. e_x^1, e_θ^1 correspond to the longitudinal and circumferential components taken, respectively, from the whole cell recording experiment and the microchamber experiment. e_x^2, e_θ^2 correspond to the longitudinal component taken independently from the two experiments.

$\gamma_a\times10^3$, N/m	1	5	10
$\Delta_x^1=\lvert(e_x^1-e_x)/e_x\rvert$	22%	19%	15%
$\Delta_\theta^1=\lvert(e_\theta^1-e_\theta)/e_\theta\rvert$	22%	15%	13%
$\Delta_x^2=\lvert(e_x^2-e_x)/e_x\rvert$	3.6%	3.6%	3.6%
$\Delta_\theta^2=\lvert(e_\theta^2-e_\theta)/e_\theta\rvert$	2.2%	2.2%	2.2%

Since the experimental conditions were different, we obtained two independent equations that allowed a re-estimation of the parameters.

We give our comparison of the re-estimated parameters with the original ones in terms of the following ratios:

$$\Delta_x^1=\left|\frac{e_x^1-e_x}{e_x}\right|, \quad \Delta_\theta^1=\left|\frac{e_\theta^1-e_\theta}{e_\theta}\right|, \tag{27}$$

$$\Delta_x^2=\left|\frac{e_x^2-e_x}{e_x}\right|, \quad \Delta_\theta^2=\left|\frac{e_\theta^2-e_\theta}{e_\theta}\right|. \tag{28}$$

The coefficients e_x, e_θ; e_x^1, e_θ^1; and e_x^2, e_θ^2 correspond, respectively, to the original estimates; the first combination where the longitudinal component is taken from the whole cell recording experiment, and the circumferential component is taken from the microchamber experiment; and the second combination where the longitudinal component is taken from two different experiments.

The difference between the original results and the re-estimates based on the second combination of the experimental data is relatively low and does not depend on γ_a (Table II). The difference is more considerable and depending on γ_a when each component of the active force is measured in a separate experiment. In the Appendix, we give an additional analysis to explain theoretically the close results obtained based on three approaches. The presented data demonstrate the consistency of our estimates with available experimental information on outer hair cell motility.

ACKNOWLEDGMENTS

The work was supported by research Grants Nos. DC02775, from National Institute on Deafness and other Communication Disorders, and KO1 AG00732, from National Institute on Aging.

APPENDIX

We use Eqs. (3) and (4) along with the following asymptotic expressions (Spector *et al.*, 1998b):

$$C_{11}/C_{22}=0.25+2.08\epsilon, \quad \epsilon=\gamma_a/C_{22}\ll1, \tag{A1}$$

$$C_{12}/C_{22}=0.5+1.08\epsilon. \tag{A2}$$

Equations (3) and (4) can be presented in the following asymptotic form:

$$-[(0.25+2.08\epsilon)\epsilon_x+(0.5+1.08\epsilon)\epsilon_\theta]=f_x/C_{22}, \quad \text{(A3)}$$

$$-[(0.5+1.08\epsilon)\epsilon_x+\epsilon_\theta]=f_\theta/C_{22}. \quad \text{(A4)}$$

Substituting the experimental values of the active strains (Dallos *et al.*, 1993), we obtain

$$0.01+0.36\epsilon=f_x/C_{22}, \quad \text{(A5)}$$

$$0.02+0.24\epsilon=f_\theta/C_{22}. \quad \text{(A6)}$$

Taking into account the volume preservation, the conditions of the whole cell recording experiment are characterized by the following equation:

$$\frac{0.5f_\theta-f_x}{C_{22}}=\epsilon^*\epsilon. \quad \text{(A7)}$$

Here ϵ^* coincides either with ϵ_x^1 when the longitudinal and circumferential components are taken from different experiments or with ϵ_x^2 when the longitudinal component is taken from two different experiments. Substituting the circumferential component of the active force from the microchamber experiment (Dallos *et al.*, 1993) and the longitudinal component from the whole cell recording experiment (Santos-Sacchi, 1989), we obtain

$$0.012+0.38\epsilon=f_x/C_{22}, \quad \text{(A8)}$$

$$0.025+0.25\epsilon=f_\theta/C_{22}. \quad \text{(A9)}$$

Similarly, with the second combination of the experimental data we have the system

$$0.01+0.37\epsilon=f_x/C_{22}, \quad \text{(A10)}$$

$$0.02+0.22\epsilon=f_\theta/C_{22}. \quad \text{(A11)}$$

The coefficients obtained in systems (A8), (A9) and (A10), (A11) are close to the corresponding coefficients of the original system (A5), (A6).

Ashmore, J. F. (**1987**). ''A fast motile response in guinea-pig outer hair cells: The cellular basis for the cochlear amplifier,'' J. Physiol. (London) **388**, 323–348.

Brownell, W. E., Bader, C. D., Bertrand, D., and de Ribaupierre, Y. (**1985**). ''Evoked mechanical responses of isolated cochlear outer hair cells,'' Science **224**, 194–196.

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.

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 **255**, 243–249.

Hallworth, R. (**1995**). ''Passive compliance and active force generation in the guinea pig outer hair cell,'' J. Neurophysiol. **74**, 2319–2328.

Hallworth, R. (**1997**). ''Modulation of OHC force generation and stiffness by agents known to affect hearing,'' 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. 524–530.

Hallworth, R., Evans, B. N., and Dallos, P. (**1993**). ''The location and mechanism of electromotility in guinea pig outer hair cells,'' J. Neurophysiol. **70**, 549–558.

Holley, M. C., and Ashmore, J. F. (**1988**). ''A cytoskeletal spring in cochlear outer hair cell,'' Nature (London) **335**, 635–637.

Iwasa, K. H. (**1994**). ''A membrane motor model for the fast motility of the outer hair cell,'' J. Acoust. Soc. Am. **94**, 2216–2224.

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 cochlea outer hair cells,'' J. Acoust. Soc. Am. **92**, 3169–3173.

Olson, E. S., and Mountain, D. C. (**1991**). ''*In vivo* measurement of basilar membrane stiffness,'' J. Acoust. Soc. Am. **89**, 1262–1275.

Russell, I. J., and Schauz, C. (**1995**). ''Salicylate ototoxicity: effects of the stiffness and electromotility of outer hair cells isolated from the guinea pig cochlea,'' Aud. Neurosci. **1**, 309–320.

Santos-Sacchi, J. (**1988**). ''Cochlear physiology,'' in *Physiology of the Ear*, edited by A. F. Jahn and J. Santos-Sacchi (Raven, New York), pp. 271–293.

Santos-Sacchi, J. (**1989**). ''Asymmetry in voltage-dependent movements of isolated outer hair cells from the organ of Corti,'' J. Neurosci. **9**, 2954–2962.

Sit, S. P., Spector, A. A., Lue, A. J.-C., Popel, A. S., and Brownell, W. E. (**1997**). ''Micropipet aspiration of the outer hair cell lateral wall,'' Biophys. J. **72**, 2812–2819.

Spector, A. A., Brownell, W. E., and Popel, A. S. (**1996**). ''A model of cochlea outer hair cell deformations in micropipette experiments: An analytical solution,'' Ann. Biomed. Eng. **24**, 241–249.

Spector, A. A., Brownell, W. E., and Popel, A. S. (**1998a**). ''Analysis of the micropipet experiment with the anisotropic outer hair cell wall,'' J. Acoust. Soc. Am. **103**, 1001–1006.

Spector, A. A., Brownell, W. E., and Popel, A. S. (**1998b**). ''Estimation of elastic moduli and bending stiffness of the anisotropic outer hair cell wall,'' J. Acoust. Soc. Am. **103**, 1007–1011.

Steele, C. R., Baker, G., Tolomeo, J., and Zetes, D. (**1993**). ''Electromechanical models of the outer hair cell,'' in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duinfuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 207–214.

Tolomeo, J. A., and Steele, C. R. (**1995**). ''Orthotropic piezoelectric properties of cochlear outer hair cell wall,'' J. Acoust. Soc. Am. **97**, 3006–3011.

Tolomeo, J. A., and Steele, C. R. (**1998**). ''A dynamic model of outer hair cell motility including intracellular and extracellular fluid viscosity,'' J. Acoust. Soc. Am. **103**, 524–534.

Ulfendahl, M., Chan, E., McConnaughey, W. B., Prost-Domansky, S., and Elson, E. (**1998**). ''Axial and transverse stiffness measures of cochlear outer hair cells suggest a common mechanical basis,'' Pflügers ArchEur. J. Physiol. **436**, 9–15.

Xue, S., Mountain, D. C., and Hubbard, A. E. (**1993**). ''Direct measurement of electrically-evoked basilar membrane motion,'' in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duinfuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 361–368.

The integration of nonsimultaneous frequency components into a single virtual pitch

V. Ciocca[a)]
Department of Speech and Hearing Sciences, University of Hong Kong, Hong Kong

C. J. Darwin
Experimental Psychology, University of Sussex, Brighton BN1 9QG, United Kingdom

(Received 12 May 1997; revised 15 April 1998; accepted 8 January 1999)

The integration of nonsimultaneous frequency components into a single virtual pitch was investigated by using a pitch matching task in which a mistuned 4th harmonic (mistuned component) produced pitch shifts in a harmonic series (12 equal-amplitude harmonics of a 155-Hz $F0$). In experiment 1, the mistuned component could either be simultaneous, stop as the target started (pre-target component), or start as the target stopped (post-target component). Pitch shifts produced by the pre-target components were significantly smaller than those obtained with simultaneous components; in the post-target condition, the size of pitch shifts did not decrease relative to the simultaneous condition. In experiment 2, a silent gap of 20, 40, 80, or 160 ms was introduced between the nonsimultaneous components and the target sound. In the pre-target condition, pitch shifts were reduced to zero for silent gaps of 80 ms or longer; by contrast, a gap of 160 ms was required to eliminate pitch shifts in the post-target condition. The third experiment tested the hypothesis that, when post-target components were presented, the processing of the pitch of the target tone started at the onset of the target, and ended at the gap duration at which pitch shifts decreased to zero. This hypothesis was confirmed by the finding that pitch shifts could not be observed when the target tone had a duration of 410 ms. Taken together, the results of these experiments show that nonsimultaneous components that occur after the onset of the target sound make a larger contribution to the virtual pitch of the target, and over a longer period, than components that precede the onset of the target sound. © *1999 Acoustical Society of America.* [S0001-4966(99)02604-1]

PACS numbers: 43.66.Ba, 43.66.Hg, 43.66.Mk [JWH]

INTRODUCTION

Several studies have demonstrated that virtual (or residue) pitch is mainly determined by the frequency of the low-numbered, resolved partials of complex tones (Moore *et al.*, 1985; Plomp, 1967; Ritsma, 1967). The results of these studies have led to the formulation of pitch perception theories which state that virtual pitch is calculated mainly on the basis of the frequencies of the resolved, low-frequency components (see, e.g., Goldstein, 1973; Terhardt *et al.*, 1982a,b). Most studies on the perception of the virtual pitch of complex sounds have employed frequency components that had simultaneous onsets and offsets (for recent reviews, see Houtsma, 1995; Hartmann, 1996). By contrast, the partials of many natural sounds start and stop at different times and do not have the same patterns of amplitude changes over time (see, e.g., Risset and Wessel, 1982). Since changes in the spectra of natural sounds may occur within periods of a few tens of milliseconds, it is likely that pitch perception processes calculate virtual pitch by integrating frequency information over time rather than by taking a snapshot of the frequency content of a complex tone at one particular moment in time.

Two types of experiment have demonstrated that pitch perception processes are sensitive to the context in which a complex sound is presented: (i) Darwin and Ciocca (1992) showed that the pitch shifts produced by mistuning the 4th harmonic (mistuned component) of a harmonic series were reduced if the mistuned component started 80 ms or more before the other harmonics; (ii) Darwin *et al.* (1995) demonstrated that a reduction in the contribution of the mistuned component to the pitch of a complex (target) sound could be obtained when a sequence of four tones, which were identical to the mistuned component, preceded the target. The effects of onset asynchrony and sequential grouping investigated by these studies showed that preceding context can be used to exclude frequency components from the calculation of virtual pitch. The current paper asks whether the virtual pitch of a harmonic series (target tone) can be affected by frequency components that stop before its onset, or start after its offset.

Hall and Peters (1981) demonstrated the temporally integrative nature of pitch perception by performing experiments on the pitch of a sequence of three 40-ms sinusoids, separated by 10-ms gaps of silence; the total duration of each sequence was 140 ms. The frequency of each sinusoid corresponded to either the 3rd, 4th, and 5th harmonic of a 200-Hz fundamental frequency (sequence 200), or the 4th, 5th, and 6th harmonic of a 180-Hz $F0$ (sequence 180). The order of the harmonics within each sequence was always from the lowest to the highest numbered. The pitch of these sequences was measured by using a discrimination task and a

a)Electronic mail: vciocca@hkusua.hku.hk

pitch matching task, and by presenting the sequences either in quiet (high S/N ratio) or in the presence of a background noise (low S/N ratio). The latter condition was included because Houtsma and Goldstein (1971; cited in Hall and Peters, 1981) had previously reported that two nonsimultaneous harmonics presented in quiet did not give rise to the perception of a single virtual pitch. Hall and Peters found that subjects heard the sequences as having a single virtual pitch, corresponding to the fundamental frequency of the partials, in the low S/N condition. In quiet, subjects heard a spectral pitch based on the frequency of the lowest and first (in order of presentation) harmonic of each sequence. These results confirmed Hall and Peter's hypothesis that pitch mechanisms must integrate frequency information over time in order to compute virtual pitch. Moreover, these findings suggest that the period over which information about a single virtual pitch is integrated (called ''pitch integration period,'' hereafter) could be as long as 140 ms.

The present study extended Hall and Peters' findings by using the pitch matching paradigm developed by Moore *et al.* (1985), and employed in more recent studies (Darwin and Ciocca, 1992; Ciocca and Darwin, 1993; Darwin *et al.*, 1994). Darwin and his colleagues measured the pitch shifts of a harmonic series (target sound) produced by mistuning the 4th harmonic (mistuned component) by various amounts. This procedure gives an estimate of the contribution of a single component to the pitch of a complex sound. The main goal of the current paper was to explore temporal order effects by measuring the pitch shifts produced by a nonsimultaneous mistuned component which could either precede or follow the presentation of the target. An additional purpose of the study was to investigate whether pitch shifts occur when a silent gap of variable duration is introduced between the target sound and a nonsimultaneous component, and whether pitch shifts occur at different silent gap durations for preceding and following components. Although a study by Rakowski and Hirsh (1980) measured the effects of the frequency of a sine wave (leading tone) on the pitch of another (target) sinusoid as a function of the duration of a silent interval which separated the two tones, to the best of our knowledge the effects of a nonsimultaneous frequency component on the virtual pitch of a complex sound have not been previously investigated.

I. EXPERIMENT 1

In this experiment, pitch shifts produced by a mistuned component which was simultaneous with the target were compared with those obtained with two types of nonsimultaneous mistuned components: (i) mistuned components which preceded the target [pre-target components; see Fig. 1(a)], and (ii) mistuned components which started as the target stopped [post-target components; see Fig. 1(b)]. There was no temporal overlap between the target and the nonsimultaneous mistuned components.

A. Method

On each trial, the subjects adjusted the pitch of a matching harmonic series to the pitch of a target complex. The

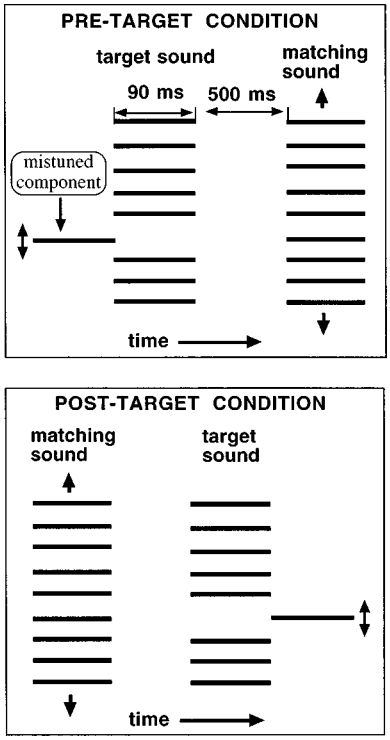


FIG. 1. Stimulus configuration for the pre-target and post-target conditions of experiment 1.

target complex contained harmonics 1–3 and 5–12 of 155 Hz, together with a possibly mistuned 4th harmonic (mistuned component). The mistuned component had a frequency of either 570, 600, 610, 620, 630, 640, or 670 Hz. These frequencies corresponded to mistunings of −8.0%, −3.2%, −1.6%, 0, +1.6%, +3.2%, and +8.0%, respectively. The matching sound consisted of the first 12 harmonics of a fundamental frequency which subjects could adjust between 151 and 159 Hz (155±4 Hz). The mistuned component, the target tone, and the matching tone had a total duration of 90 ms, including 5-ms rise/fall raised-cosine ramps. The duration of the stimuli was the same as the duration used in previous pitch matching experiments (Darwin and Ciocca, 1992; Ciocca and Darwin, 1993), and was selected because it was found to give reliable pitch shift estimates while being short enough for the purpose of the current study (see also experiment 3, Sec. III).

Different matching tasks, administered in separate sessions, were used to measure pitch shifts for pre- and post-target components. In the pre-target task, the onset-to-onset time between the mistuned component and the target was either 0 (simultaneous components) or 90 ms [pre-target components; see Fig. 1(a)], and the matching tone followed the target sound. In the post-target task, mistuned components could either start at the same time as the target (simultaneous components) or start 90 ms after the onset of the target [i.e., they target as the target stopped; see Fig. 1(b)]; the order of presentation of the target and matching tones was the opposite to that of the pre-target condition (i.e., the matching tone always preceded the target tone). The change in the order of the stimuli for performing the matching task with post-target components was introduced to ensure that

the mistuned component affected only the pitch of the target tone. If the same matching task (i.e., if the matching tone followed the target tone) had been used to measure pitch shifts for post-target components, the latter would have intervened between the target and the matching tones. In this case, mistuned components might have affected the pitch of both the matching and the target tones, or could have acted as distractors, thereby making the matching task generally more difficult. Informal observations indicated that adjusting the pitch of the matching tone to match the pitch of a subsequent target tone proved to be slightly more difficult than matching the pitch of the tones in the standard order of presentation. However, the increase in the variability of pitch matches was not very large, and listeners learned to perform the post-target matching task as accurately as the pre-target task over the course of an experimental session (see Sec. I B). In both tasks, a 500-ms silent interval separated target and matching tones.

All components were presented to the left ear at the same amplitude (corresponding to 58 dB, sound-pressure level, for a 1000-Hz tone). All components started at sine (0) phase. Sounds were presented through Sennheiser HD414 headphones in a double-walled IAC booth. At the beginning of each trial, the fundamental frequency of the matching tone was chosen at random from the permitted range (155±4 Hz). Subjects adjusted the frequency of the matching sound by moving a roller-ball up or down (Darwin and Ciocca, 1992). Subjects were allowed to perform as many adjustments as they required in order to achieve a satisfactory match. Each target sound was matched five times in a quasi-random order within each experimental session; one session took 1–2 h to complete, including rest breaks. Within a trial, the mistuned components, especially a nonsimultaneous one, could often be heard out as an individual pure tone. Subjects were instructed to ignore the mistuned component and to focus their attention on matching the pitches of the two complex sounds.

To summarize, the experiment consisted of 21 conditions: 7 frequencies (570, 600, 610, 620, 630, 640, or 670 Hz) by 3 types of mistuned component (simultaneous, pre-target, or post-target). The two nonsimultaneous conditions were run in separate sessions since they required different matching tasks. In each session, the simultaneous and the nonsimultaneous conditions were repeated five times. Pre- and post-target sessions were run on separate days, and their order was counterbalanced across subjects such that half the subject participated in the pre-target session first, and vice versa.

Sounds were synthesized in real-time at 44.1 kHz using custom software (Russell and Darwin, 1991) written for the 56001 processor of the Digidesign Audiomedia II board, and output through that board's 16-bit DACs and anti-aliasing filters. The board was attached to an Apple MacIIcx computer which controlled the experiment.

Eight university students, six of whom were musically trained, participated in the experiment. All had participated in previous pitch matching experiments.

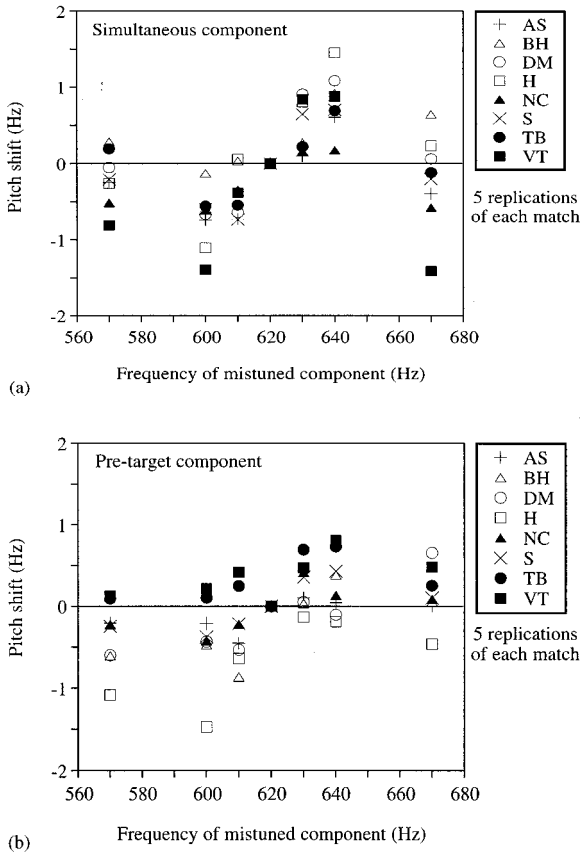


FIG. 2. Individual pitch shifts, averaged over five repetitions, for the pre-target task of experiment 1. Pitch shifts for the simultaneous components are shown in (a); shifts for the pre-target mistuned components are displayed in (b).

B. Results and discussion

The effect of the mistuning of the 4th harmonic on the pitch of the target tone was determined by measuring the difference between the pitch matches for each amount of mistuning relative to the condition in which the 4th harmonic was not mistuned. The negative and positive pitch shifts for simultaneous and pre-target components in the pre-target task are displayed in Fig. 2(a) and (b), respectively. Pitch shifts for individual listeners were calculated as the difference between the average of pitch matches for each −ve/+ve mistuning and the average of the pitch matches for the in-tune, 620-Hz condition (0-Hz point on the *y*-axis). The pattern of pitch shifts for the simultaneous components replicates the pattern obtained with similar stimuli in previous pitch matching experiments; pitch shifts were largest for mistunings of ±20 Hz, and were on average close to 0 Hz for mistunings of ±50 Hz (see, for example, Darwin and Ciocca, 1992). The pitch shifts for pre-target components were generally smaller than those produced by simultaneous components, but there were individual differences. Subjects showed a large reduction in the size of pitch shifts either for negative mistunings (subjects VT, NC, and TB), or for positive mistunings of pre-target components (subjects DM, BH, and H) relative to the simultaneous component condition. The remaining subjects (AS and S) showed a reduction of pre-target shifts for both negative and positive mistunings [see Fig. 2(b)].

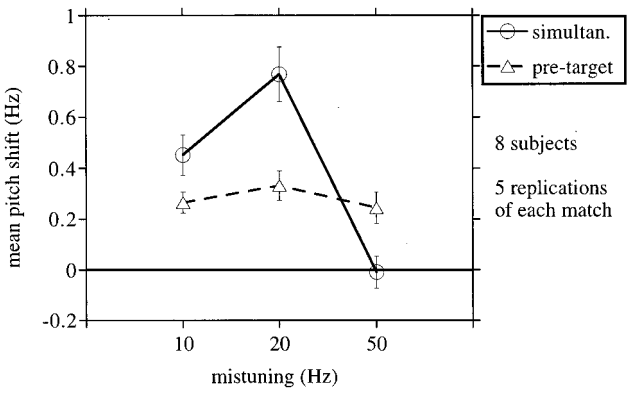


FIG. 3. Mean pitch shifts with error bars (standard errors of the mean) for the pre-target task of experiment 1. Mean shifts for the pre-target mistuned component are displayed as open triangles, and those for the simultaneous component as open circles.

The mean pitch shifts for the pre-target task, averaged across eight subjects, are displayed in Fig. 3. Mean pitch shifts were calculated as half the difference between matches for each amount of positive and negative mistuning for each subject (Moore *et al.*, 1985). A two-way ANOVA with repeated measures was applied to the individual mean pitch shifts obtained in the pre-target sessions. The factors were the ''mistuning'' of the mistuned component (10, 20, or 50 Hz), and the ''type of mistuned component'' (simultaneous or pre-target). The main effect of ''mistuning'' was statistically significant [$F(2,14)=19.07$, $p<0.0001$]. In the simultaneous condition, mean pitch shifts were largest at mistunings of 20 Hz (3.2% of the harmonic frequency), and were virtually absent at mistunings of 50 Hz. Shifts for both the 10- and 20-Hz mistunings were significantly higher than for mistunings of 50 Hz (Tukey HSD test, $p<0.01$). The effects of mistuning with simultaneous components is consistent with previous findings which showed that pitch shifts are at a maximum for mistunings of about 3%, and then decrease to zero with mistunings of 8% (Darwin and Ciocca, 1992; Moore *et al.*, 1985). In the pre-target condition, only the 20-Hz mistuning produced shifts which were significantly different from those of the 50-Hz, simultaneous condition (Tukey HSD test, $p<0.05$). Pitch shifts for the three pre-target mistunings did not differ from each other (Tukey HSD test, $p>0.05$). The difference in the effect of mistuning between the simultaneous and the pre-target conditions resulted in a significant ''mistuning'' by ''type of mistuned component'' interaction [$F(2,14)=14.44$, $p<0.0005$]. The main effect of ''type of mistuned component'' showed that overall, pitch shifts were smaller in the pre-target than in the simultaneous conditions [$F(1,7)=5.59$, $p<0.05$]. Pitch shifts for the simultaneous components were significantly larger than those for pre-target components only for mistunings of 20 Hz (Tukey HSD test, $p<0.01$); for mistunings of 10 and 50 Hz, pitch shifts were not significantly different (Tukey HSD test, $p>0.05$).

The negative and positive pitch shifts for the simultaneous and the post-target components in the post-target task are displayed in Fig. 4(a) and (b), respectively. The overall pattern and size of pitch shifts for the post-target and for the

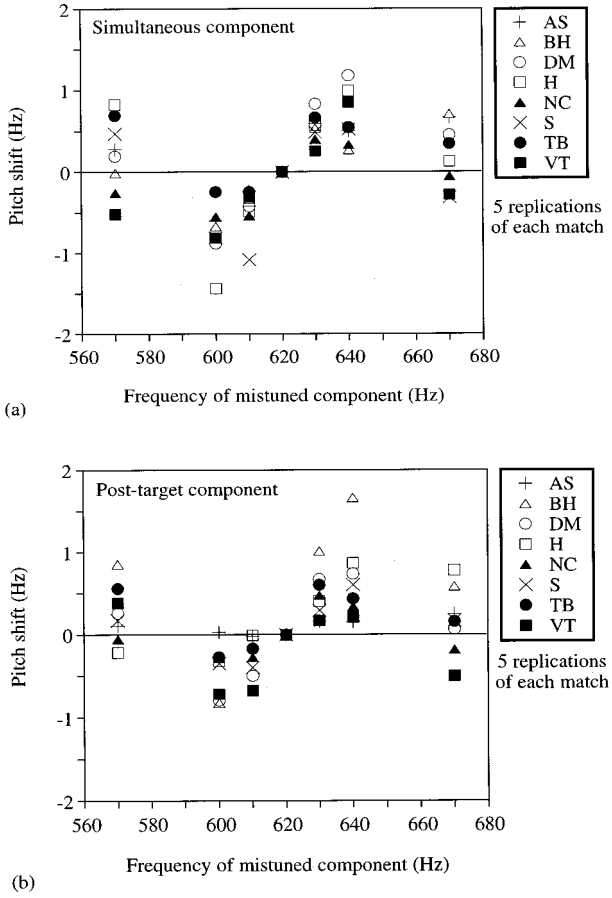


FIG. 4. Individual pitch shifts, averaged over five repetitions for the post-target task of experiment 1. Pitch shifts for the simultaneous components are shown in (a); shifts for the post-target mistuned components are displayed in (b).

simultaneous components is similar. For most subjects (six out of eight) a reduction in the size of the pitch shifts occurred either for positive or for negative mistunings of 10–20 Hz in the post-target condition [Fig. 4(b)] relative to the simultaneous condition [Fig. 4(a)]; for two listeners, BH and AS, the patterns of pitch shifts for post-target components differed from that of the other subjects. Listener AS, who showed virtually no pitch shifts in the post-target condition, also showed very small shifts in the pre-target condition. This listener was perhaps able to make a reliable estimate of the pitch of the target very quickly, possibly by focusing only on the first few tens of milliseconds of the target sound. Unlike AS, listener BH perceived much larger pitch shifts in the post-target than in the simultaneous condition for positive mistunings.

A two-way ANOVA with repeated measures was applied to the mean pitch shifts (averages of $-$ve and $+$ve pitch shifts for each amount of mistuning) for the post-target session (see Fig. 5). The factors were ''mistuning'' (10, 20, or 50 Hz) and ''type of mistuned component'' (simultaneous or post-target). The main effect of ''mistuning'' was statistically significant [$F(2,14)=36.25$, $p<0.0001$]. The pattern of pitch shifts as a function of mistuning is very similar in the simultaneous and the post-target conditions: Pitch shifts were largest at 20-Hz mistunings and decreased to about zero at mistunings of 50 Hz. The effect of ''type of mistuned

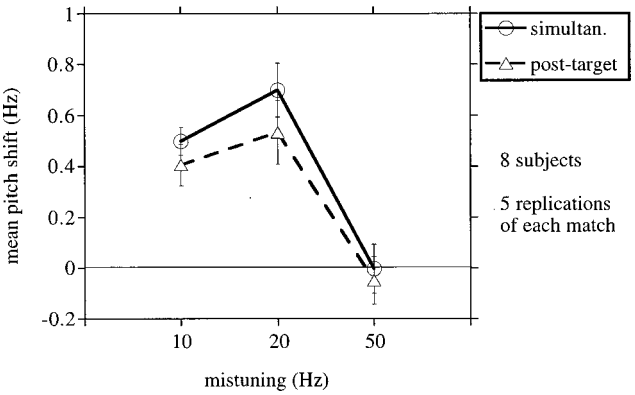


FIG. 5. Mean pitch shifts with error bars (standard errors of the mean) for the post-target task of experiment 1. Mean shifts for the post-target mistuned component are displayed as open triangles, and those for the simultaneous component as open circles.

component'' was not statistically significant [$F(1,7)=2.96$, $p>0.05$]. The difference between simultaneous and post-target components was not statistically significant for any amount of mistuning (Tukey HSD tests). The ''mistuning'' by ''type of mistuned component'' interaction was not statistically significant [$F(2,14)=0.15$, $p>0.05$].

In order to perform a direct comparison of the effects of a preceding and a following nonsimultaneous component on the pitch of the target sound, a two-way ANOVA with repeated measures was applied to the mean pitch shifts for the pre- and post-target conditions. The factors were the ''mistuning'' (10, 20, or 50 Hz) and the ''type of mistuned component'' (pre-target or post-target). Pre- and post-target mistuned components had different effects on the mean pitch shifts as a function of the amount of mistuning, as shown by a statistically significant ''type'' by ''mistuning'' interaction [$F(2,14)=5.62$, $p<0.05$]. The main effect of ''mistuning'' was also statistically significant [$F(2,14)=11.93$, $p<0.001$]; the main effect of ''type'' was not statistically significant [$F(1,7)=0.21$, $p>0.05$].

To summarize, this experiment showed that a mistuned component which is nonsimultaneous can produce pitch shifts in a target complex. This finding extends the results of Hall and Peters' (1981) study which found that listeners integrated a sequence of three brief sinusoids into a single virtual pitch when these stimuli were presented in noise (low S/N ratio condition). While Hall and Peters reported a lack of virtual pitch perception in the quiet condition of their study, the integration of the nonsimultaneous component into the virtual pitch of the target tone was observed in quiet in the present experiment. This difference could be due to the nature of the stimuli: When only three nonsimultaneous harmonics are used, low S/N ratios may be necessary to be able to perceive a single virtual pitch. On the other hand, the virtual pitch of the target complex could be perceived in quiet, and the current task measured whether virtual pitch was affected by the presence of a nonsimultaneous component (instead of measuring whether a virtual pitch could be perceived at all). Second, this experiment showed that pitch shifts in a target complex sound were smaller when the mistuned component preceded the target than when the target and the mistuned component were presented simultaneously. By contrast, the post-target and simultaneous conditions produced virtually identical pitch shifts. The fact that the contribution of a nonsimultaneous component to the pitch of the target complex is asymmetrical with respect to the order of presentation of the stimuli might indicate that pitch perception mechanisms calculate virtual pitch in a ''forward'' fashion. That is, pitch processes integrate acoustic information which follows the onset of the complex into a single virtual pitch; acoustic energy which precedes the onset of a harmonic complex would be less strongly integrated into the pitch of the complex. This asymmetry may be explained by the fact that the onset of a harmonic complex causes pitch perception mechanisms to search for components which could ''fit'' into the pitch of a complex sound. This search could be based on a ''harmonic sieve'' principle (Scheffers, 1983). A post-target component would therefore contribute to the pitch of the target complex as long as its frequency is not too far from the harmonic frequency. When the mistuned component precedes the target complex (pre-target condition), the calculation of the virtual pitch of the target sound would not begin until immediately after the offset of the component; therefore, the mistuned component would be less likely to be included in the pitch of the harmonic complex.

II. EXPERIMENT 2

If the pitch processor calculates virtual pitch in a forward fashion then the pitch integration period should be relatively long following the onset of the harmonic complex (post-target components), but it should be short for information that precedes the onset of the complex sound (pre-target components). This prediction was tested in the second experiment by varying the duration of a silent gap which separated pre-target and post-target mistuned components from the target tone.

A. Method

The pre- and post-target tasks were run in separate experimental sessions as in the first experiment. In both tasks, the duration of the silent interval which separated the target from the nonsimultaneous mistuned component was varied and could be either 0, 20, 40, 80, or 160 ms. In the pre-target task, the interval separated the offset of the mistuned component from the onset of the target; in the post-target task, the silent interval separated the offset of the target complex and the onset of the mistuned component. A condition in which the mistuned component and the target tone were simultaneous was also included in both experimental session. Experiment 1 showed that mistunings of $\pm 3\%$ (600- and 640-Hz mistuned components) produced the largest pitch shifts in the target sound, in agreement with the findings of previous pitch matching studies (see, e.g., Darwin and Ciocca, 1992). Therefore, only three mistuned components (600, 620, and 640 Hz) were included in this experiment.

The stimuli were played through Sennheiser HD 250 Linear earphones in a single walled IAC booth which was located in a sound-insulated room. The software/hardware

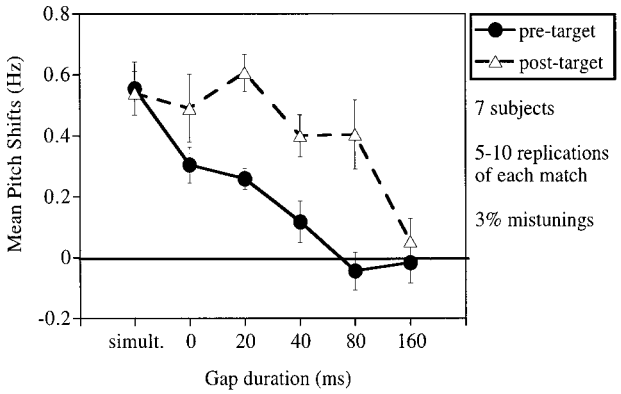


FIG. 6. Mean pitch shifts with error bars (standard errors of the mean) for the pre- and post-target tasks of experiment 2. Mean shifts for mistunings of ±3% are displayed for the simultaneous component and for each interval duration of the pre-target (filled symbols) and post-target tasks (open symbols).

setup was the same as in the previous experiment, except that a PowerMac 7100/70AV was used to run the experiment instead of a MacIIcx.

Seven listeners, one of whom was the first author, took part in the experiment; three of the listeners were musically trained. All listeners had participated in previous pitch matching experiments, but none had taken part in experiment 1. Subjects completed between five and ten matches for each experimental condition, depending on their pitch matching accuracy after five matches per condition. The criterion for stopping after five matches was a standard error smaller than 0.4 Hz in all conditions. The experiment was run in separate blocks during which subjects completed 2–3 matches for each condition; 2–4 blocks were run by each listener, with each block taking 1–2 h to complete.

B. Results

Mean pitch shifts for the pre- and post-target sessions, averaged across listeners, are displayed in Fig. 6. A two-way ANOVA with repeated measures was applied to the mean pitch shifts for each subject in each experimental condition. The ''temporal overlap'' between the mistuned component and the target sound (simultaneous, 0-, 20-, 40-, 80-, or 160-ms gap), and the ''type of mistuned component'' (pre-target versus post-target) were the experimental factors. Mean pitch shifts were larger for post-target than for pre-target components (main effect of ''type'') [$F(1,6)=12.26$, $p<0.05$]. The overall pattern of shifts indicates that pitch shifts decreased as the silent gap duration increased, as shown by a statistically significant main effect of ''temporal overlap'' [$F(5,30)=16.18$, $p<0.0001$].

For pre-target components, the 0-ms condition produced lower mean pitch shifts than the simultaneous component (planned comparison, $p<0.02$); this finding replicates the results obtained in the 20-Hz mistuning condition of experiment 1. Pitch shifts were effectively eliminated when a silent gap of 80 ms separated the offset of the mistuned component and the onset of the target; this condition was significantly different from the 20-ms condition (planned comparison, $p<0.005$). The 40- and 80-ms conditions were not significantly different from each other (planned comparison, $p>0.05$). The results of the post-target task replicated those of experiment 1, which showed that the pitch shifts for the synchronous and the 0-ms delay conditions were virtually identical for mistunings of 20 Hz. Mean pitch shifts dropped to zero in the 160-ms gap condition, which was significantly different from the 0-ms condition (planned comparison, $p<0.001$). The 40- and 80-ms conditions were not significantly different from the 0-ms condition (planned comparison, $p>0.05$).

As the delay between the nonsimultaneous component and the target was increased, mean pitch shifts decreased at different rates for pre-target and for post-target components (''temporal overlap'' by ''type'' interaction) [$F(5,30)=3.11$, $p<0.05$]. The mean pitch shifts for pre- and post-target components were significantly different in the 20-ms, 40-ms, and 80-ms conditions (planned comparisons, $p<0.01$). The pre- versus post-target difference in pitch shifts for the 0-ms condition just failed to reach statistical significance (planned comparison, $p=0.07$).

To summarize, pitch shifts were significantly reduced when the silent interval between the pre-target component and the target complex was 20 ms or longer; when a silent gap of 80 ms (or longer) was introduced between the offset of the component and the onset of the target, pitch shifts were virtually eliminated. For post-target components, pitch shifts were identical to those produced by a simultaneous component for delays of up to 40 ms; pitch shifts were reduced to zero only for the longest delay (160 ms). These results support the previous suggestion that the pitch integration period extends for a longer interval following than preceding the onset of a complex sound.

III. EXPERIMENT 3

The results of the previous experiments have been considered to be indicative of the duration of the ''pitch integration period'' (i.e., the temporal window within which pitch processes integrate acoustic information into a single virtual pitch) with respect to the onset of the target. This interpretation is based on the assumption that, in the current task, the formation of virtual pitch (i) is triggered by the onset of the target, (ii) includes a relatively short interval prior to the onset of the target, and (iii) ends once a reliable estimate of virtual pitch has been achieved following the onset of the target. If this assumption is valid, then pitch shifts obtained as a function of silent gap duration would represent an estimate of the pitch integration period. However, an alternative interpretation of the results is that pitch processes might not cease the calculation of virtual pitch as soon as a reliable estimate of the pitch of the target has been made; instead, pitch processes may keep revising their estimate as long as information is received. Consider, for example, the post-target stimuli: it is possible that listeners use only the last portion of the target in conjunction with the nonsimultaneous component for performing the pitch matching task. If this were the case, the inclusion of the duration of the target into the estimate of the duration of the pitch integration period would not be warranted. A simple test of these hypotheses

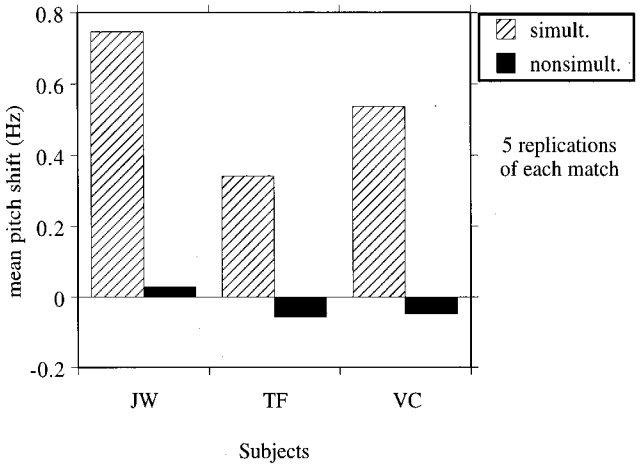


FIG. 7. Mean pitch shifts for the stimuli of experiment 3. Mean shifts for mistunings of ±3% are displayed for the simultaneous and the post-target conditions.

can be carried out by using stimulus durations in excess of the longest estimated duration of the pitch integration period (i.e., about 200 ms for post-target stimuli).

This experiment employed the post-target task only, since post-target components produced pitch shifts for longer gap durations than pre-target components in the previous experiments. The stimuli had a fixed duration of 410 ms; this duration was the same as that of the stimuli used in previous pitch matching tasks in which pitch shifts were measured for simultaneous mistuned components (Darwin and Ciocca, 1992). Only the post-target condition with no silent gap between the target and the nonsimultaneous component (0-ms delay) was tested. If subjects perceive the pitch of the target sound by integrating acoustic information from the onset of the target until a reliable pitch estimate has been achieved then nonsimultaneous components should not produce pitch shifts in this experiment, since the pitch of the target would be calculated well before the offset of the target. By contrast, if the alternative hypothesis is correct then we would expect to observe pitch shifts for 410-ms target sounds followed by nonsimultaneous components of the same duration.

A. Method

The stimuli were identical to those used in the post-target session of experiment 1, except that: (i) all the stimuli had a fixed duration of 410 ms instead of 90 ms, and (ii) only mistunings of ±20 Hz were used. Therefore, a total of six experimental conditions were employed in this experiment: two types of mistuned component (simultaneous versus post-target) by three frequencies (600, 620, and 640 Hz). Subjects completed five matches for each condition.

The apparatus was identical to that employed in experiment 2.

Three listeners, one of whom was the first author, took part in the experiment. All listeners had taken part in previous pitch matching experiments, and two had taken part in experiment 2; one of the listeners was musically trained. The experiment took about 30 min to complete.

B. Results

The pitch shifts for −ve and +ve mistunings, averaged across five trials, are displayed in Fig. 7 for each subject. This figure shows that the nonsimultaneous mistuned components produced no pitch shifts, while the simultaneous components did. Since the pattern of pitch shifts for simultaneous and nonsimultaneous components was identical for the three subjects, it was not considered necessary to perform any statistical analysis on the data. The results of this experiment fully support the idea that listeners begin calculation of the virtual pitch at the onset of the target tone, and complete such calculation as soon as a reliable pitch estimate has been achieved. This process is completed before the offset of the stimuli used in the present experiment, which suggests that the pitch integration period for the current stimuli is shorter than 410 ms. This conclusion supports the interpretation of the nonsimultaneous pitch shifts obtained in experiment 2 in terms of estimated duration of the pitch integration period.

IV. GENERAL DISCUSSION

A. Summary

The present study investigated the integration of a nonsimultaneous mistuned component into the virtual pitch of a harmonic complex. The first experiment showed that, for brief target tones: (i) the pitch shifts produced by a post-target mistuned component were virtually identical to the shifts produced by a simultaneous mistuned component; (ii) pre-target components produced significantly smaller pitch shifts than simultaneous components. The findings of experiment 2 indicated that pitch shifts occurred at longer gap durations for post-target components (up to 80 ms) than for pre-target components (up to 40 ms). The third experiment demonstrated that the current pitch matching procedure can be employed for obtaining reliable estimates of the duration of the pitch integration period with respect to the onset of the target sound.

B. Matching virtual or spectral pitches?

The interpretation of the results of this study in terms of the contribution of a nonsimultaneous component to the virtual pitch of the target tone relies on the assumption that subjects were matching the virtual pitch of the target and matching tones. It has been suggested that in pitch discrimination or matching experiments, subjects could perform these tasks by comparing the spectral pitches of corresponding frequency components instead of comparing the virtual pitches of the complex tones (Faulkner, 1985). In principle, the matching of the spectral pitch of the mistuned component with the corresponding 4th harmonic of the (fully harmonic) matching tone might well have occurred in the present study, particularly when the component was nonsimultaneous and could be easily heard as having a distinct spectral pitch. However, there are several reasons for rejecting the idea that listeners were matching spectral pitches of individual components in the present experiments. First, Moore (1987) demonstrated that pitch shifts produced by (simultaneous) mistuned components can be measured even when the

matching tone does not contain the harmonic that corresponded to the mistuned component. Second, if listeners matched the pitch of the mistuned component, pitch shifts should have been larger than those normally observed in pitch matching studies. For example, in the present experiments pitch shifts were about ± 1 Hz, on average, when the 4th harmonic was mistuned by ± 20 Hz. If subjects had adjusted the matching tone so that its 4th harmonic matched the pitch of the mistuned component, then pitch shifts should have been much larger than those observed in the present study. Moreover, the matching of the spectral pitch of mistuned components would have produced a monotonic increase in the size of pitch shifts with increasing mistuning,[1] which was not the case. Third, matching the spectral pitch of the mistuned component should have been easier, thereby producing larger pitch shifts, in the nonsimultaneous than in the simultaneous conditions. Finally, matching the pitch of the mistuned component should have been easier in the pre-target conditions, in which the mistuned component was the first sound within a trial, than in the post-target conditions, in which the first sound within a trial was the matching tone having a randomly selected fundamental frequency. Contrary to these predictions, pitch shifts were never larger in the nonsimultaneous than in the simultaneous conditions, and were usually larger in the post- than in the pre-target condition. These considerations support the idea that, in the present study, subjects matched the virtual pitches of the target and matching tones rather than matching the spectral pitches of individual components.

C. Temporal integration and pitch perception

Although further research is needed to obtain a precise estimate of the duration and of the window shape of the pitch integration period, a tentative model of temporal pitch integration can be formulated from the present data by assuming that: (i) the temporal integration window is rectangular; (ii) the calculation of virtual pitch is triggered by the onset of the target; (iii) the pitch integration period includes a relatively short interval preceding, and a longer interval following, the onset of the target; (iv) the endpoint of the pitch integration period occurs within the interval spanning from the longest gap duration at which pitch shifts were observed to the gap duration at which pitch shifts were eliminated. While (i) is assumed for simplicity, criteria (ii), (iii), and (iv) can be justified on the basis of the results of experiments 2 and 3. It might be argued that criterion (ii) applies to post-target but not to pre-target conditions, because subjects might have started the calculation of the virtual pitch of the target as soon as they heard a pre-target component at the beginning of a trial (in spite of instructions to ignore pre-target components and focus on the matching of the pitch of target and matching sounds). However, this hypothesis does not explain the decrease in pitch shifts with increasing delays in the pre-target conditions of experiment 2, nor is it clear how the pitch of the target could be calculated on the basis of a single mistuned component. It is more likely that pitch estimation is initiated at the onset of the target, but that the pitch integration period includes an interval prior to the onset of the target, as specified by criterion (iii). This criterion is consistent with the assumptions that the auditory input is first stored into a ''preperceptual auditory image,'' and that auditory processes such as those involved in pitch perception operate on the information stored in the preperceptual auditory image (Massaro, 1972).

On the basis of these criteria, the duration of the pitch integration period can be estimated to be 40–80 ms prior to target onset, and 170–250 ms following the onset of the target sound (the latter duration also includes the target tone duration of 90 ms). Robinson and Patterson (1995) investigated the identification of the pitch chroma of vowels as a function of the number of cycles, and found that performance reached an asymptote at 16–32 cycles, which corresponds to durations of 122–244 ms and 92–184 ms for $F0$'s of 131 and 174 Hz, respectively. Their results are consistent with the results of experiment 3, which suggest that pitch perception processes integrate acoustic information into a single virtual pitch until a reliable estimate has been generated. The current estimates are also in agreement with the results of studies which investigated fundamental frequency discrimination (Carlyon, 1996; Plack and Carlyon, 1995). For example, Carlyon (1996) demonstrated that listeners perceptually integrate portions of a masker which occurred before and after a 200-ms target complex into the target's virtual pitch over a period of 100 ms or longer (''overintegration''). Carlyon found that overintegration occurred when the target was composed of unresolved, but not of resolved, harmonics. However, it is possible that overintegration could be observed with resolved harmonics for target durations shorter than 200 ms. Evidence in support of this prediction is found in two recent studies. Plack and Carlyon (1995) reported that the $F0$ discrimination of the fundamental frequency of two 50-ms complex tones deteriorated when duration of the gap between the tones was reduced from 500 ms to 50 ms, even when the two tones consisted of resolved harmonics. Micheyl and Carlyon (1998) found that $F0$ discrimination of 100-ms target tones composed by resolved components was degraded by the presence of complex tones which preceded and followed the targets. Further evidence for the fact that nonsimultaneous resolved components can be integrated into a single pitch comes from studies which employed pure tone stimuli (Rakowski and Hirsh, 1980; Kelly and Watson, 1986).

The current estimate of the duration of the pitch integration period is consistent with previous suggestions that the auditory system integrates acoustic energy over a period of few hundreds of milliseconds for the perception of loudness (''loudness summation;'' Zwislocki, 1969), backward and forward masking (Wilson and Carhart, 1971), and amplitude modulation detection (Sheft and Yost, 1990). The present results are also compatible with temporal integration models based on shorter time constants (few milliseconds), such as the ''multiple look'' model proposed by Viemeister and Wakefield (1991). This model states that the auditory system takes samples or ''looks'' of the incoming acoustic signal about every 3 ms. These samples are stored in memory and can be selected for further processing. The present data could be explained by such model if it is assumed that pitch processes can combine information across frequency channels

from samples which occurred over periods up to about 200 ms, despite the temporal separation between the nonsimultaneous component and the other components of the target tone (see, for example, White and Carlyon, 1997, for evidence in support of this kind of temporal integration).

The present findings suggest that pitch processes calculate virtual pitch by giving a higher weighting to acoustic information which follows the onset of a complex sound than to information which precedes it. This temporal asymmetry complements the finding that onset asynchrony is more effective than offset asynchrony in preventing an asynchronous harmonic from contributing to the timbre of a vowel (Darwin, 1984). This asymmetry is also consistent with the results of previous studies which showed that the final portion of a signal is given a higher weighting than initial information for the perception of the pitch of frequency glides (Nabelek *et al.*, 1970) and ramped versus damped sinusoids (Patterson, 1994a,b). A larger degradation in performance with trailing maskers (which followed the presentation of target tones) than with leading maskers was observed in order discrimination (Kelly and Watson, 1986), and pitch identification tasks (''backward recognition masking;'' Massaro, 1975). However, the effects of trailing maskers in those studies were not frequency specific, and were interpreted in terms of the interference of the maskers with the stored representation of (brief) target sounds. Therefore, it is not clear that this interference effect played a major role in the pitch matching paradigm used in the present study, which measured the integration of pre- and post-target components into the virtual pitch of the target.

D. Implications for models of pitch perception

The current findings, which imply that pitch processes integrate acoustic energy over a period of few hundred milliseconds for calculating virtual pitch, have implications for models of pitch perception. First, these findings suggest that phase information is not likely to be an important factor for the integration of a frequency component into a single virtual pitch, since mistuned components which followed the offset of the target by up to 80 ms of silence produced relatively large pitch shifts. It is interesting to note that, unlike pitch perception, the detection of the mistuning of a harmonic was found to be affected by its phase (Hartmann, 1988). Second, these results are difficult to reconcile with models of pitch perception that estimate virtual pitch by combining autocorrelation functions of the activity of individual nerve fibers (see recent models by Meddis and Hewitt, 1991a,b; Yost *et al.*, 1996; Meddis and O'Mard, 1997). These models calculate virtual pitch by summing the output of autocorrelation functions across frequency channels; the summary autocorrelation function thus obtained (''summary autocorrelogram'') is then used to estimate virtual pitch, which corresponds to the period of the highest peak in the summary autocorrelogram. The models require that the energy of frequency components which is used to produce a summary autocorrelogram occur within a ''time constant,'' which is typically a period of few milliseconds (see also a related model by de Cheveigne, 1998). The results reported here, as well as those of the studies reviewed above, have demonstrated that the integration of nonsimultaneous acoustic energy for pitch perception can occur over periods of 100 ms or longer. In order to account for these findings, an autocorrelation model would have to be able to store autocorrelation functions for each frequency channel and for periods of at least 100 ms, and then apply some temporal smoothing before the summary autocorrelogram is calculated.[1] The integration of nonsimultaneous energy over relatively long durations for the perception of virtual pitch could be perhaps more easily implemented by models which compute virtual pitch from the frequency of resolved components (Goldstein, 1973; Terhardt *et al.*, 1982a). These models could be modified to specify differential weightings of frequency components for the calculation of pitch, depending on the components' time of occurrence with respect to the onset of complex sounds.

ACKNOWLEDGMENTS

We would like to thank: Robert Carlyon and Chris Plack for providing helpful comments on the manuscript; Paul Russell for software support; Duncan Lam, Chloe Lam, and Ida Tsui for helping with data collection. Experiment 1 was funded through Grant No. S.E.R.C. GR/F 34060 awarded to the second author, and grant British Council UK/HK JRS 93/8 awarded to both authors. Experiments 2 and 3 were funded through Grant No. H.K.R.G.C. HKU 362/94M awarded to the first author.

[1]We are grateful to Robert Carlyon for this suggestion.

Carlyon, R. P. (**1996**). ''Masker asynchrony impairs the fundamental-frequency discrimination of unresolved harmonics,'' J. Acoust. Soc. Am. **99**, 525–533.

Ciocca, V., and Darwin, C. J. (**1993**). ''Effects of onset asynchrony on pitch perception: Adaptation or grouping?,'' J. Acoust. Soc. Am. **93**, 2870–2878.

Darwin, C. J. (**1984**). ''Perceiving vowels in the presence of another sound: constraints on formant perception,'' J. Acoust. Soc. Am. **76**, 1636–1647.

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. **9**, 3381–3390.

Darwin, C. J., Ciocca, V., and Sandell, G. J. (**1994**). ''Effects of frequency and amplitude modulation on the pitch of a complex tone with a mistuned harmonic,'' J. Acoust. Soc. Am. **95**, 2631–2636.

Darwin, C. J., Hukin, R. W., and Al-Khatib, B. Y. (**1995**). ''Grouping in pitch perception: Evidence for sequential constraints,'' J. Acoust. Soc. Am. **98**, 880–885.

de Cheveigne, A. (**1998**). ''Cancellation model of pitch perception,'' J. Acoust. Soc. Am. **103**, 1261–1271.

Faulkner, A. (**1985**). ''Pitch discrimination of harmonic complex signals: Residue pitch or multiple component discriminations,'' J. Acoust. Soc. Am. **78**, 1993–2004.

Goldstein, J. L. (**1973**). ''An optimum processor theory for the central formation of the pitch of complex tones,'' J. Acoust. Soc. Am. 1496–1516.

Hall, J. W., and Peters, R. W. (**1981**). ''Pitch from nonsimultaneous successive harmonics in quiet and noise,'' J. Acoust. Soc. Am. **69**, 509–513.

Hartmann, W. M. (**1988**). ''Pitch perception and the segregation and integration of auditory entities,'' in *Auditory Function—Neurobiological Bases of Hearing*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowen (Wiley, New York), pp. 623–645.

Hartmann, W. M. (**1996**). ''Pitch, periodicity, and auditory organization,'' J. Acoust. Soc. Am. **100**, 3491–3502.

Houtsma, A. J. M. (**1995**). ''Pitch perception,'' in *Hearing*, edited by B. C. J. Moore (Academic, San Diego), pp. 267–295.

Houtsma, A. J. M., and Goldstein, J. L. (**1971**). *Perception of musical intervals: Evidence for the central origin of the pitch of complex* (Tech. Re. 484): MIT, Cambridge.
Kelly, W. J., and Watson, C. S. (**1986**). ''Stimulus-based limitations on the discrimination between different temporal orders of tones,'' J. Acoust. Soc. Am. **79**, 1934–1938.
Massaro, D. W. (**1972**). ''Stimulus information vs processing time in auditory pattern recognition,'' Percept. Psychophys. **12**, 50–56.
Massaro, D. W. (**1975**). ''Backward recognition masking,'' J. Acoust. Soc. Am. **58**, 1059–1065.
Meddis, R., and Hewitt, M. (**1991a**). ''Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification,'' J. Acoust. Soc. Am. **89**, 2866–2882.
Meddis, R., and Hewitt, M. (**1991b**). ''Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II: Phase sensitivity,'' J. Acoust. Soc. Am. **89**, 2883–2894.
Meddis, R., and O'Mard, L. (**1997**). ''A unitary model of pitch perception,'' J. Acoust. Soc. Am. **102**, 1811–1820.
Micheyl, C., and Carlyon, R. P. (**1998**). ''Effects of temporal fringes on fundamental frequency discrimination,'' J. Acoust. Soc. Am. **104**, 3006–3018.
Moore, B. C. J. (**1987**). ''The perception of inharmonic complex tones,'' in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 180–189.
Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (**1985**). ''Relative dominance of individual partials in determining the pitch of complex tones,'' J. Acoust. Soc. Am. **77**, 1853–1860.
Nabelek, I. V., Nabelek, A. K., and Hirsh, I. J. (**1970**). ''Pitch of tone bursts of changing frequency,'' J. Acoust. Soc. Am. **48**, 536–553.
Patterson, R. D. (**1994a**). ''The sound of a sinusoid: spectral models,'' J. Acoust. Soc. Am. **96**, 1409–1418.
Patterson, R. D. (**1994b**). ''The sound of a sinusoid: time-domain models,'' J. Acoust. Soc. Am. **96**, 1419–1428.
Plack, C. J., and Carlyon, R. P. (**1995**). ''Differences in frequency modulation detection and fundamental frequency discrimination between complex tones consisting of resolved and unresolved harmonics,'' J. Acoust. Soc. Am. **98**, 1355–1364.
Plomp, R. (**1967**). ''Pitch of complex tones,'' J. Acoust. Soc. Am. **41**, 1526–1533.
Rakowski, A., and Hirsh, I. J. (**1980**). ''Post-stimulatory pitch shifts for pure tones,'' J. Acoust. Soc. Am. **68**, 467–474.
Risset, J. C., and Wessel, D. L. (**1982**). ''Exploration of timbre by analysis and synthesis,'' in *The Psychology of Music*, edited by D. Deutsch (Academic, New York), pp. 25–58.
Ritsma, R. J. (**1967**). ''Frequencies dominant in the perception of the pitch of complex sounds,'' J. Acoust. Soc. Am. **42**, 191–198.
Robinson, K., and Patterson, R. D. (**1995**). ''The stimulus duration required to identify vowels, their octave, and their pitch chroma,'' J. Acoust. Soc. Am. **98**, 1858–1865.
Russell, P., and Darwin, C. J. (**1991**). ''Real-time synthesis of complex sounds on a Mac II with 56001 DSP chip,'' Br. J. Audiol. **25**, 59–60.
Scheffers, M. T. (**1983**). ''Sifting vowels: Auditory pitch analysis and sound segregation,'' Unpublished Ph.D. thesis, Groningen University, The Netherlands.
Sheft, S., and Yost, W. A. (**1990**). ''Temporal integration in amplitude modulation detection,'' J. Acoust. Soc. Am. **88**, 796–805.
Terhardt, E., Stoll, G., and Seewann, M. (**1982a**). ''Algorithm for extraction of pitch salience from complex tonal signals,'' J. Acoust. Soc. Am. **71**, 679–688.
Terhardt, E., Stoll, G., and Seewann, M. (**1982b**). ''Pitch of complex signals according to virtual pitch theory,'' J. Acoust. Soc. Am. **71**, 671–678.
Viemeister, N. F., and Wakefield, G. H. (**1991**). ''Temporal integration and multiple looks,'' J. Acoust. Soc. Am. **90**, 858–865.
White, L. J., and Carlyon, R. P. (**1997**). ''Detection of signals having expected and unexpected temporal structures,'' Hearing Res. **112**, 141–146.
Wilson, R. H., and Carhart, R. (**1971**). ''Forward and backward masking: Interactions and additivity,'' J. Acoust. Soc. Am. **49**, 1254–1263.
Yost, W. A., Patterson, R., and Sheft, S. (**1996**). ''A time domain description for the pitch strength of iterated rippled noise,'' J. Acoust. Soc. Am. **99**, 1066–1078.
Zwislocki, J. J. (**1969**). ''Temporal summation of loudness: An Analysis,'' J. Acoust. Soc. Am. **46**, 431–441.

Monosyllabic word recognition at higher-than-normal speech and noise levels

Gerald A. Studebaker[a] and Robert L. Sherbecoe
Memphis Speech & Hearing Center, The University of Memphis, 807 Jefferson Avenue, Memphis, Tennessee 38105

D. Michael McDaniel
Communication Disorders, Arkansas State University, P.O. Box 910, State University, Arkansas 72467

Catherine A. Gwaltney
Memphis Speech & Hearing Center, The University of Memphis, 807 Jefferson Avenue, Memphis, Tennessee 38105

(Received 27 February 1998; revised 17 June 1998; accepted 18 December 1998)

The effects of intensity on monosyllabic word recognition were studied in adults with normal hearing and mild-to-moderate sensorineural hearing loss. The stimuli were bandlimited NU#6 word lists presented in quiet and talker-spectrum-matched noise. Speech levels ranged from 64 to 99 dB SPL and S/N ratios from 28 to −4 dB. In quiet, the performance of normal-hearing subjects remained essentially constant; in noise, at a fixed S/N ratio, it decreased as a linear function of speech level. Hearing-impaired subjects performed like normal-hearing subjects tested in noise when the data were corrected for the effects of audibility loss. From these and other results, it was concluded that: (1) speech intelligibility in noise decreases when speech levels exceed 69 dB SPL and the S/N ratio remains constant; (2) the effects of speech and noise level are synergistic; (3) the deterioration in intelligibility can be modeled as a relative increase in the effective masking level; (4) normal-hearing and hearing-impaired subjects are affected similarly by increased signal level when differences in speech audibility are considered; (5) the negative effects of increasing speech and noise levels on speech recognition are similar for all adult subjects, at least up to 80 years; and (6) the effective dynamic range of speech may be larger than the commonly assumed value of 30 dB. © *1999 Acoustical Society of America.* [S0001-4966(99)00904-2]

PACS numbers: 43.66.Ba, 43.66.Sr, 43.66.Ts, 43.71.Ky [JWH]

INTRODUCTION

There has long been evidence that the speech recognition performance of normal-hearing listeners deteriorates when speech is transmitted at very high intensities, i.e., at levels above about 90 dB SPL. Fletcher may have been the first to consider this topic when, in 1922, he reported that the intelligibility of nonsense syllables presented in quiet decreased when the syllables were highly amplified (Fletcher, 1922). Later papers by French and Steinberg (1947) and Fletcher and Galt (1950) reconfirmed this early finding, while studies by Kryter (1946), Pickett and Pollack (1958), and Pollack and Pickett (1958) provided additional information about the negative effects of amplification when speech was mixed with noise. In each of these latter experiments, it was found that monosyllabic word scores dropped significantly when speech at a constant S/N ratio was increased from 80 to 130 dB SPL.

Some research has also suggested that reductions in speech recognition performance can occur when signals are presented at SPLs well below those used in the studies cited above. However, the literature offers contradictory information about this possibility. For example, studies by Hawkins and Stevens (1950), Hirsh and Bowman (1953), Hirsh *et al.* (1954), and Duquesnoy and Plomp (1983) all indicate the S/N ratio needed for 50%-correct performance remains constant at speech levels through at least 90 dB SPL. In contrast, the results of various other studies indicate performance begins to decrease as soon as speech at a constant S/N ratio exceeds conversational intensity levels (Speaks *et al.*, 1967; Chung and Mack, 1979; Dirks *et al.*, 1982; Hagerman, 1982; Beattie, 1989; Goshorn and Studebaker, 1994).

People with hearing losses often listen to higher-than-normal levels of speech and noise when they wear amplification devices. In theory, most of these people should benefit from added amplification because it raises the sensation level of the speech signal. This increases audibility and, in turn, intelligibility. However, when amplifier gain controls are turned up, speech and noise spectral levels increase together, with the result that the S/N ratio remains constant within any given band; thus, intelligibility may or may not increase as much as a simple gain model would suggest. Further, if hearing-impaired subjects who listen under conditions of constant audibility experience the same negative effects of increased level as the normal-hearing subjects tested in some studies, then turning up the gain control could actually reduce speech intelligibility for the hearing impaired.

A survey of various studies that have obtained performance-intensity (P-I) functions for speech materials presented at moderate-to-high levels suggests the following, sometimes conflicting, conclusions. First, in the case of both

[a]Electronic mail: gstudbkr@memphis.edu

normal-hearing subjects and hearing-impaired subjects, speech recognition scores decrease when speech becomes uncomfortably loud (Dirks *et al.*, 1981, 1982). Second, some hearing-impaired subjects show lower scores and larger reductions in score with increasing level than normal-hearing subjects, while others do not (Jerger and Jerger, 1971; Dirks *et al.*, 1981). Third, the S/N ratio needed to maintain a constant level of performance remains constant over level for both hearing-impaired and normal-hearing subjects (Duquesnoy and Plomp, 1983). Fourth, the difference in performance between average normal-hearing and average hearing-impaired subjects decreases with increasing level (Dirks *et al.*, 1982). Fifth, when the S/N ratio remains constant, scores for normal-hearing subjects begin to decrease at speech levels approaching or exceeding those of conversation (Hagerman, 1982; Goshorn and Studebaker, 1994), while scores for hearing-impaired subjects do not begin falling until much higher levels are presented (Dirks *et al.*, 1981).

Unfortunately, for the reconciliation of these findings, the roles and interactions of absolute signal level, changing audibility, and hearing status are often unclear. In most cases, performance increases with level when audibility is free to vary (e.g., French and Steinberg, 1947; Fletcher and Galt, 1950; ANSI, 1997), and decreases with level when audibility remains fixed (Kryter, 1946; Pickett and Pollack, 1958; Hagerman, 1982). Typically, however, such audibility effects have not been adequately taken into account. Thus, data on the relationship between speech level and performance, especially for the hearing impaired, are often confounded by the positive and negative consequences of increased level.

To understand how hearing impairment influences speech recognition at higher than normal listening levels, it is necessary to evaluate these offsetting mechanisms separately. One way to do this is to hold audibility constant while varying the speech or noise level (Pickett and Pollack, 1958; Dirks *et al.*, 1982; Hagerman, 1982). This is a method commonly used with normal-hearing subjects. However, it may not be practical with many hearing-impaired subjects because it is difficult or impossible to keep audibility constant over speech level when the subjects have more than a mild hearing loss. Another method is to test normal-hearing subjects in a masking noise that causes their thresholds to match those of hearing-impaired subjects (Fabry and Van Tasell, 1986; Humes *et al.*, 1987; Zurek and Delhorne, 1987; Dubno and Schaefer, 1992; Hargus and Gordon-Salant, 1995). This method too is limited by the magnitude of the hearing losses that can be emulated. In addition, it does not keep audibility constant as speech level is varied. It just equalizes the change in audibility for the hearing-impaired and masked-normal-hearing subjects. Thus, only relative, i.e., between-group, comparisons are possible. The absolute effects of increased level cannot be assessed in either group.

A third approach that might be used, either alone or in combination with these two methods, is to calculate and correct for any audibility changes that occur as speech level is varied. If reasonable estimates of these audibility effects could be obtained, a wider range of hearing losses potentially could be evaluated and the comparisons between normal-hearing and hearing-impaired subjects could be improved.

The purpose of the present study was to determine how listening level and S/N ratio influence speech recognition by normal-hearing and hearing-impaired subjects at higher-than-normal speech and noise levels. We were also curious whether a very simple articulation index (AI) style calculation could be used to compare hearing-impaired subjects with normal-hearing subjects when audibility cannot be controlled experimentally. To answer this question, performance was tested as a function of speech level in quiet and noise at a series of fixed S/N ratios. The results were analyzed both with and without corrections for audibility effects.

I. METHODS

A. Subjects

Three groups of subjects were tested. Group NH consisted of 72 normal-hearing young adults (23 males, 49 females). Group HI-1 consisted of 32 hearing-impaired adults (22 males, 10 females) under age 70 years. Group HI-2 consisted of 12 hearing-impaired adults (7 males, 5 females) age 70 years or older.

The subjects in group NH were college undergraduates recruited from the student populations of the University of Memphis (U-of-M) and Arkansas State University (ASU). They had pure-tone thresholds of 15 dB HL or better at the octave frequencies from 250 to 8000 Hz (ANSI, 1989) and acoustic immittance results that were consistent with normal middle-ear function.

The subjects in groups HI-1 and HI-2 were adults who had previously received routine audiological evaluations at the U-of-M Speech and Hearing Center or the ASU Speech and Hearing Clinic. Only people with long-standing bilateral sensorineural hearing losses, no obvious signs of middle-ear disease, and pure-tone averages (500, 1000, 2000 Hz) greater than 25 dB HL were selected. Twenty-two of the subjects in group HI-1 (69%) and ten of the subjects in group HI-2 (83%) were hearing aid wearers.

Table I reports age and hearing-threshold data for all three subject groups. The threshold data are based on clinical audiometric measurements obtained using a step size of 5 dB.

B. Test conditions

The test conditions consisted of bandlimited speech presented in quiet and in a slightly wider band of filtered masking noise. The signal levels and filter bandwidths were chosen based on the results of a pilot study at the U-of-M involving 16 normal-hearing adults. None of the subjects in that study participated in the current experiment.

The speech materials were copies of the Auditec of St. Louis recordings of the NU#6 word test, spoken by a male talker. They were created by digitizing a set of commercially prepared audiotapes and storing the output on removable Bernoulli disk cartridges (Studebaker *et al.*, 1993). The masker was random noise that had been digitally filtered so

TABLE I. Age and hearing threshold statistics for each subject group. The table reports means, ranges, and standard deviations (s.d.) for age (in years) and pure-tone hearing level (rounded to the nearest decibel) at the octave frequencies (Hz).

Group	Statistic	Age	250	500	1000	2000	4000	8000
NH	Mean	22.9	5	5	3	5	3	5
	Range	18–36	−5–15	−5–15	−5–15	−5–15	−10–15	−10–15
	s.d.	3.5	4	4	4	5	5	6
HI-1	Mean	59.3	23	25	33	50	63	62
	Range	40–69	0–55	10–60	15–65	25–75	35–110	25–90
	s.d.	8.3	13	13	14	11	17	18
HI-2	Mean	75	29	32	38	51	67	78
	Range	70–80	10–50	15–55	25–65	35–80	35–90	55–100
	s.d.	3.8	14	11	11	12	17	12

that its long-term rms spectrum, measured in 50-Hz-wide bands, matched the 1% levels of the NU#6 talker's short-term (120-ms) speech spectrum.

The speech passband (447–2239 Hz) was limited to the most important frequencies for NU#6 word recognition, i.e., the 1/3-octave bands with centers ranging from 500 to 2000 Hz (Studebaker *et al.*, 1993). The noise passband (282–2818 Hz) covered the entire speech band, including the filter skirts.

These frequency limits were chosen based on the following three considerations. First, we wanted to avoid scores of 100% in quiet and 0% at the poorest S/N ratios. Second, we did not want to include frequencies near 4000 Hz because some hearing-impaired subjects have sharply falling threshold curves that would have made the planned corrections for audibility effects more difficult. Third, we wanted to keep the noise levels as low as possible.

A wider speech band, more like conversational speech, might have revealed greater masking and/or level effects. However, the chosen bandwidth encompasses most of the frequencies needed for speech understanding, and is not unlike that experienced by many hearing aid wearers. In addition, both the literature (Rosenthal *et al.*, 1975; Milner *et al.*, 1984; Goshorn and Studebaker, 1994) and our own calculations suggested that a wide bandwidth would have only a small, and therefore negligible, effect on performance under the conditions of this study.

A total of eight speech long-term rms levels (64, 69, 74, 79, 84, 89, 94, 99 dB SPL) and ten S/N ratios (−4, −1, 2, 5, 8, 12, 16, 20, 24, 28 dB) were evaluated. The speech levels were 1 dB below the SPLs obtained in quiet with a wideband filter (110–11 000 Hz). The noise levels varied, depending on the S/N ratio presented. The S/N ratios listed in this paper are the values that would be produced if the speech and noise signals had the same bandwidth and infinitely sharp filter skirts. The measured S/N ratios were approximately 1 dB less than the reported values because the actual noise bandwidth and skirts were slightly wider than those for the speech.

C. Equipment

The testing system consisted of a personal computer with an internal Bernoulli cartridge drive (Iomega, 44 mbyte) and a 12-bit D/A converter to play the speech (Data Translation, DT-2821), an attenuator (Hewlett Packard, 350D) to limit the output from the converter, a noise source, separate filter channels for the speech and noise, and an audiometer (Madsen, OB-822). Each filter channel was composed of a preprocessing analog filter (speech: Wavetek, 751A; noise: Wavetek, 852), a digital filter (IHR Universal filter, UDF-III) and a postprocessing analog filter (Wavetek, 852). This setup provided about 200 dB/oct of attenuation in the low frequencies and more than 1900 dB/oct of attenuation in the high frequencies.

The filtered speech and noise signals were adjusted to desired levels and combined using the audiometer; then they were passed to a single-insert earphone (Etymotic Research, ER-3A) placed in the subject's ear. All testing was performed with the subject seated inside a double-walled sound-treated room. The nontest ear was not masked or plugged.

With one exception, the same equipment was used to evaluate all three subject groups. The exception was that in the U-of-M tests on normal-hearing subjects, the masking noise was produced by playing white noise from a noise generator (General Radio, 1382) through the filters and into the audiometer. In contrast, in all the other tests, the noise came from a recording of the filter output that was played directly into the audiometer using a digital audiotape player (Tascam DA-30). Measurements obtained with a dynamic signal analyzer (Hewlett Packard, HP-3561A) indicated that the two noise sources produced virtually the same spectrum at the earphone.

D. Calibration

Equipment calibration was performed at the start of the experiment and on a daily basis, before each test session. Initially, the unfiltered NU#6 test words, digitally joined together without intervening silent periods, were used to calibrate the speech channel; the shaped noise signal was used to calibrate the noise channel. Daily calibration was carried out using a 1000-Hz pure tone with a known relationship to the rms levels of these signals.

Speech and noise measurements were collected using a Zwislocki coupler (Knowles Electronics, DB-100) and a Type I sound level meter (Larson-Davis, 800-B) set for *C* scale, integrate mode, and a 3-dB exchange rate. Coupler

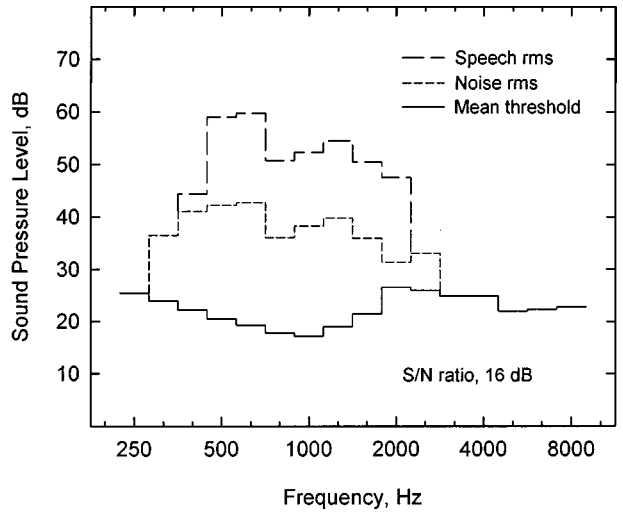


FIG. 1. 1/3-oct band spectra (in dB, sound-pressure level) for the speech and noise signals and for the mean hypothesized internal noise levels of the normal-hearing subjects. The levels reflect long-term rms values and include the frequency effects of the earphone and other equipment used in common at the two test sites.

levels, in dB SPL (2 min L_{eq}), were read directly from the meter's LCD screen. When pure-tone measurements were obtained, the sound level meter was set for slow, linear weighting, and 1/3-oct mode.

Figure 1 displays representative speech and noise spectra for a typical listening condition (speech level of 64 dB SPL, 16 dB S/N ratio). The spectra were measured using the signal analyzer's digitally synthesized 1/3-oct band filters and include the effects of the insert earphone and other equipment common to the two testing sites. The long-dash line shows the long-term rms speech spectrum; the short-dash line shows the rms noise spectrum; and the solid line shows the mean hypothesized internal noise spectrum for the normal-hearing subjects based on their audiometric thresholds for pure tones.[1]

E. Procedures

Initially, the subjects received one or two unfiltered practice lists, in quiet at 65 dB SPL. Digitized recordings of the CID W-22 word test (Auditec of St. Louis) were used for this purpose. Then, the experimental conditions involving NU#6 word lists were presented, in randomized order without replacement.

To reduce possible subject biases that might result from repeated exposure to the same speech materials, the test conditions were divided into blocks of 16 conditions each, and a given subject was allowed to receive only one of the possible blocks. Table II reports the condition block presented to each group of listeners. The conditions for the hearing-impaired groups were selected after all the normal-hearing subjects had been tested and the difficulty of the various blocks could be estimated.

The normal-hearing subjects received all 16 conditions in one session, while the hearing-impaired subjects received eight conditions on one day and eight on another day. These different testing schedules were used because many of the hearing-impaired subjects were elderly and we did not want their performance to be biased by fatigue. The mean time between the test sessions given to the hearing-impaired subjects was about 6 days for group HI-1 and about 4 days for group HI-2.

Every subject, whether normal hearing or hearing impaired, completed one 50-item NU#6 word list under each condition. The order of the four lists was rotated across the subjects and conditions so that the different lists were presented an equal number of times. The test ear was chosen randomly for the normal-hearing subjects. Each hearing-impaired subject was tested using his or her better ear, based on measurements of the subject's pure-tone average HL for 500, 1000, and 2000 Hz.

Computer software developed at the University of Memphis (Matesich, 1991) was used to control the testing process. The subjects sat in front of a computer terminal that displayed the item numbers and either typed their answers on a keyboard (normal-hearing subjects) or wrote them by hand on preprinted test forms (hearing-impaired subjects). Normal-hearing subjects used the keyboard's enter key to control the presentation rate of the stimuli; hearing-impaired subjects pressed a hand-held electronic switch. At the end of the experiment, the subjects were paid for their participation.

TABLE II. Experimental conditions presented to each subject group. The table reports the test site (U-of-M =University of Memphis, ASU=Arkansas State University), the number of listeners evaluated per condition (N), the long-term rms speech levels (dB SPL) presented, and the S/N ratios (dB) used.

Group	Site	N	Speech levels				S/N ratios			
NH	U-of-M	12	64	74	84	94	Quiet	12	5	−1
NH	U-of-M	12	69	79	89	99	Quiet	12	5	−1
NH	U-of-M	12	64	74	84	94	16	8	2	−4
NH	U-of-M	12	69	79	89	99	16	8	2	−4
NH	ASU	12	64	74	84	94	28	24	20	16
NH	ASU	12	69	79	89	99	28	24	20	16
HI-1	U-of-M	16	69	79	89	99	Quiet	12	5	−1
HI-1	ASU	16	69	79	89	99	Quiet	12	5	−1
HI-2	U-of-M	6	69	79	89	99	Quiet	12	5	−1
HI-2	ASU	6	69	79	89	99	Quiet	12	5	−1

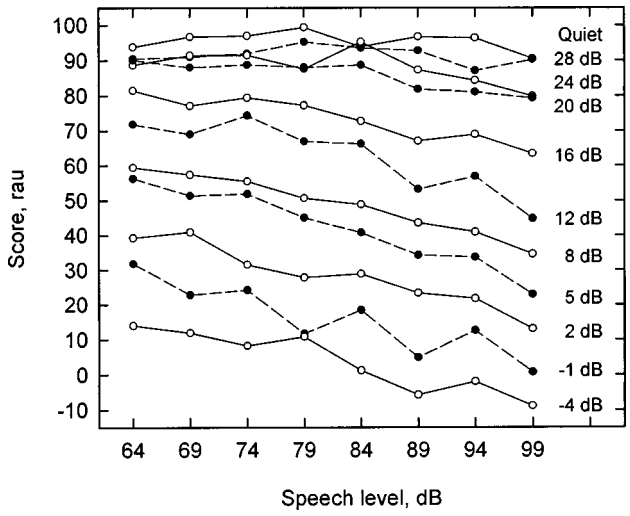


FIG. 2. Mean word-recognition scores (in rau) for the normal-hearing subjects, plotted as a function of speech sound-pressure level.

F. Data treatment

Answers obtained from the normal-hearing subjects were scored with a computerized scoring program (Matesich, 1993). Answers obtained from the hearing-impaired subjects were scored manually by two of the experimenters. All results were transformed into rationalized arcsine units (rau) before any statistical analyses so that the variance would be homogeneous over the full range of the scores obtained (Studebaker, 1985).

Transformed scores for the normal-hearing subjects were averaged across subjects and test sites.[2] Scores for the hearing-impaired subjects were first evaluated by age group. Then, the results for the two age groups were combined. The average results were compared with those for the normal-hearing subjects on an absolute basis and after correction for the calculated effects of audibility change.

II. RESULTS

A. Normal-hearing subjects

1. Performance versus speech level

Figure 2 shows mean scores for the normal-hearing subjects, plotted as a function of speech level. Each function displays the results at a different and constant S/N ratio. These functions are hereafter referred to as PICA (performance intensity at constant audibility) functions to differentiate them from traditional performance-intensity (P-I) functions where S/N ratio is varied by holding either just the speech or just the noise level constant.

The figure reveals that performance declined substantially as the speech and noise levels increased. An interaction between the effects of level and S/N ratio is also apparent. Reductions in the mean score were noticeably larger when the S/N ratio was below 16 dB than when it was above that point. However, the reductions at higher S/N ratios may be more noteworthy because they suggest that the range of speech levels that contribute to intelligibility may be larger than the 30 dB normally assumed in calculation procedures such as the AI. The reasons for this are discussed later.

TABLE III. Equations and constants used to derive the best-fit straight lines shown in Fig. 3. Equation (1) was used to determine the slope (in rau/dB) of each best-fit line. Equation (2) was used to determine the intercept (in rau) of each best-fit line.

Eq. (1): $D = a\ \exp(-0.5((x-b)/c)^2)$.

(a)	(b)	(c)
−0.8244	4.2599	11.8561

Eq. (2): $I = \overline{Y} - D\overline{X}$.

S/N ratio (dB)	Slope (D)	Intercept (I)
Quiet (33)	−0.0437	99.13
28	−0.1110	100.61
24	−0.2061	105.04
20	−0.3415	113.53
16	−0.5049	114.54
12	−0.6662	117.17
8	−0.7844	112.75
5	−0.8228	109.06
2	−0.8096	94.27
−1	−0.7471	76.78
−4	−0.6468	56.42

At most S/N ratios, the mean scores display a strong linear relationship to speech level. Possible exceptions are evident at the highest S/N ratios, and in quiet, where the PICA functions are modestly curved. Even in these cases, however, the discrepancies are not very large. It was therefore decided that straight lines would be used to evaluate the mean data. These lines were fit to the data as follows.

First, straight lines were independently fit to the PICA functions at each S/N ratio using traditional least-squares regression. The slopes of these independent lines were then plotted versus S/N ratio and analyzed using nonlinear regression. The results indicated that the variations in slope were well described by a Gaussian function having the form shown in Eq. (1).[3]

$$D = a\ \exp(-0.5((x-b)/c)^2). \tag{1}$$

In this equation, D refers to the function density (the slope value) at S/N x, and a, b, and c, are fitting constants representing the maximum slope, the S/N ratio associated with that slope, and the standard deviation of the function, respectively. The values of a, b, and c that produced the best fit are reported in Table III (df adjusted $r^2 = 0.9225$, fit std error $= 0.0706$ rau). The values of D obtained for each of the S/N ratios used in this study are also reported in Table III.

The next step was the derivation of the origin I for each line having a slope defined by Eq. (1). This was done using Eq. (2) where X represents the mean speech level and Y is the mean score value obtained at a given S/N ratio.

$$I = \overline{Y} - D\overline{X}. \tag{2}$$

The values of I obtained for each of the S/N ratios are reported in Table III.

Finally, the slope and intercept values provided by Eqs. (1) and (2) were used to plot straight lines for each S/N. These are shown in Fig. 3, along with the raw data from Fig. 2. A line for the quiet condition is also included. This line, which is described more fully in the next section, assumes

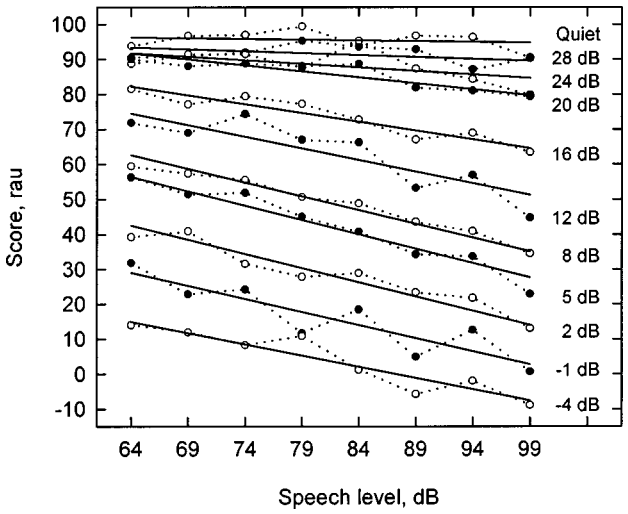


FIG. 3. Straight lines fit to the mean word-recognition scores (in rau) from Fig. 2.

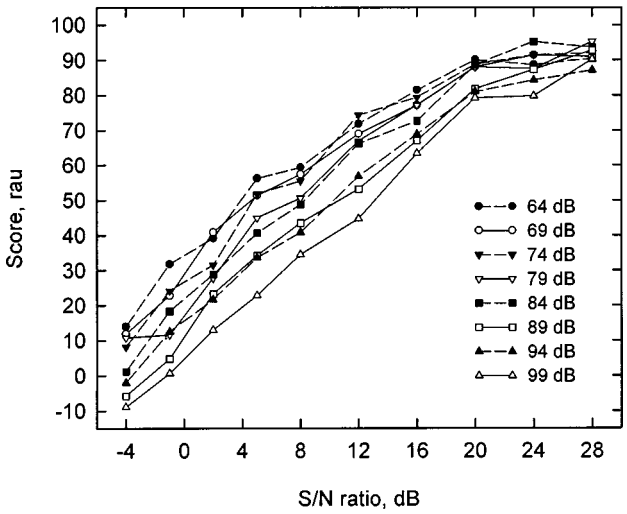


FIG. 4. Mean word-recognition scores (in rau) for the normal-hearing subjects, plotted as a function of signal-to-noise (S/N) ratio.

that speech presented in quiet is equivalent to speech presented at an S/N ratio of 33 dB. The overall rms difference between the lines and the data was 2.87 rau.

2. Speech self-masking

In their classic 1947 publication, French and Steinberg hypothesized that what they called the ''equivalent noise of self-masking'' was 24 dB below the long-term average speech level. Operationally, this self-masking level is the same as the low-level boundary of the speech signal's dynamic range.

The data from the present study provide a basis for estimating the equivalent noise of self-masking. The estimate was made as follows. First, Eq. (1) and the coefficients from Table III were used to predict a slope (D) for every S/N ratio at 1-dB intervals between 28 and 36 dB. Next, the corresponding Y-intercept (I) was determined for each of these lines using Eq. (2), as described earlier. Then, the scores predicted by each line were compared to the scores obtained in quiet. The smallest rms difference between the predicted and observed data was obtained when a S/N ratio of 33 dB was assumed. This value can be compared directly to the 24 dB value of French and Steinberg, and it is the source of the S/N ratio associated with the ''quiet'' data in Fig. 3, mentioned earlier.

3. Performance versus S/N ratio

To evaluate the results of this study in a more conventional way, the mean scores obtained at each speech level were replotted as a function of S/N ratio; that is, as P-I functions. The results are shown in Fig. 4. The separation of the different P-I functions reveals the effects of overall speech and noise level on performance. The functions also show that these effects occurred over a broad range of S/N ratios. The P-I function slopes for every speech level reveal that the mean score increased at a roughly constant rate until an S/N ratio of 12 dB was reached. Slopes in this region ranged from 3.4 to 4.0 rau/dB and had a mean value of 3.7 rau/dB.

Between 12 and 20 dB S/N, the functions for the lower speech levels become less steep while the functions for the higher speech levels maintain nearly the same slope. Possible reasons for this outcome include: performance ceiling effects, an unequal distribution of difficult words, and reduced importance in the low-intensity part of the speech signal. Given the nature of this particular experiment, it is not possible to rule out any of these causes, but performance ceiling effects seem less likely than the others given the modest scores obtained and the use of the rau transform.

Above 20 dB S/N, scores for all the speech levels continue to increase but at a much slower rate, averaging only about 0.7 rau/dB. This dramatic reduction in P-I function slope could also be due to the factors cited earlier. Again, however, it is not possible to rule out any of these possible contributors.

B. Hearing-impaired subjects

1. Subject age

Figure 5 reports mean word-recognition scores for the younger versus the older hearing-impaired subject group, plotted as a function of speech level. Notice that the results for the two groups do not differ appreciably except in the quiet condition and at the lowest speech level with an S/N of 12 dB. Calculations suggested that these outcomes are readily explained by the fact that the older subjects had slightly poorer pure-tone thresholds than the younger subjects within the speech passband (see Table I). Because of this, we concluded that audibility rather than age was responsible for any apparent differences in the performance of the two hearing-impaired groups, and that their data could be combined.

2. Performance versus speech level

Figure 6 shows the pooled results for the hearing-impaired subjects, plotted as a function of speech level. As in the case of the normal-hearing subjects, the functions describe performance when S/N ratio is held constant. Notice,

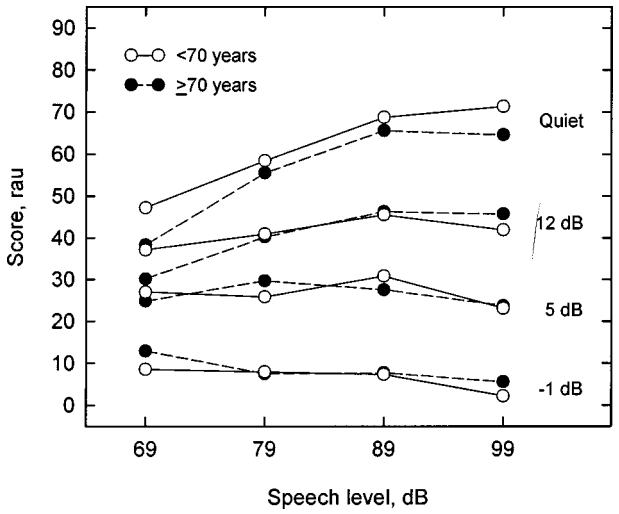


FIG. 5. Mean word-recognition scores (in rau) for the two hearing-impaired subject groups, plotted as a function of speech sound-pressure level.

however, that the performance of these subjects either did not decrease or else decreased only slightly. In fact, in quiet and at the best S/N ratio, their mean performance actually increased.

A plausible explanation for such results is that they reflect the combined effects, both positive and negative, of increasing level. When the speech was presented in quiet to the hearing-impaired subjects, it was often below threshold, while at the better S/N ratios, the masking noise was often below threshold. Thus, whenever the speech level increased under these conditions, speech audibility improved and performance increased. On the other hand, at the poorer S/N ratios, audibility was largely determined by the noise and did not improve, or improved only slightly, when the speech and noise levels were increased. Therefore, performance at the poorer S/N ratios either remained constant or declined with increasing level, as in the case of the normal-hearing subjects.

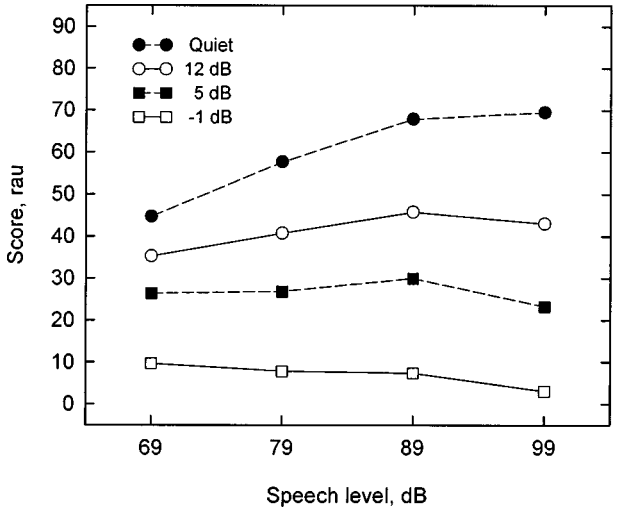


FIG. 6. Mean word-recognition scores (in rau) for all of the hearing-impaired subjects, plotted as a function of speech sound-pressure level.

3. Audibility correction

To test this hypothesis, the mean results for the hearing-impaired subjects were corrected for changes in audibility using a simple AI-type procedure. The resulting corrected functions were then compared with a set of similarly corrected functions derived for the normal-hearing subjects. Clearly, the validity of this approach depends upon the validity of the method used to adjust the data for differences in audibility. Therefore, a calculation procedure that made as few assumptions as possible was used.

The first step in this procedure was the determination of an audibility index value for each test condition. This was accomplished as follows. An S/N ratio was calculated for each 1/3-octave band between 125 and 10 000 Hz by taking the difference between the speech level and either the applied noise level or the subjects' average internal noise level, whichever was higher. Fifteen dB were added to this S/N ratio to convert it into a speech peaks-to-noise ratio (Steeneken and Houtgast, 1980; ANSI, 1997). Each speech-peak-to-noise ratio was then divided by 40, the assumed speech dynamic range, and multiplied by the appropriate frequency importance weight for the Auditec NU#6 speech test (Studebaker *et al.*, 1993). The audibility index was equal to the sum of all the 1/3-octave-band products produced in this manner.

The value for the dynamic range of speech was based on the normal-hearing subject data obtained under the condition where the speech was at 69 dB SPL. It was derived by iteratively varying the assumed range until the mean and the standard deviation of the differences between the predicted and observed scores at this speech level were as close to zero as possible. This process resulted in two ''optimal'' estimates for the dynamic range (39 and 41 dB) that were then averaged to obtain the final value (40 dB). Refer to Sec. III F of this paper for a more detailed discussion of this value and its effect on the appearance of the data. The use of a single constant as a divisor presumes that the range does not vary with frequency and that every decibel within the range has equal weight. The same value was used in all the audibility index calculations.

The next step was to convert each index value into a predicted score using Eq. (3), which is based on a formula reported by Fletcher and Galt (1950).

$$S=(1-10^{-A/Q})^{N}. \tag{3}$$

Here, S is the word score expressed as a proportion; A is the audibility index; and Q and N are fitting constants. This equation was selected because previous studies have shown that it provides a good description of the relationship between index values and monosyllabic word scores (Studebaker and Sherbecoe, 1991; Studebaker *et al.*, 1993). Values for the fitting constants Q and N were taken from a study on the intensity importance function for speech (Studebaker *et al.*, 1997). Although these values differed somewhat from the published coefficients for the NU#6 recordings (Studebaker *et al.*, 1993), they produced virtually identical predicted results. They were used in preference to the original values only because they were derived with a procedure that

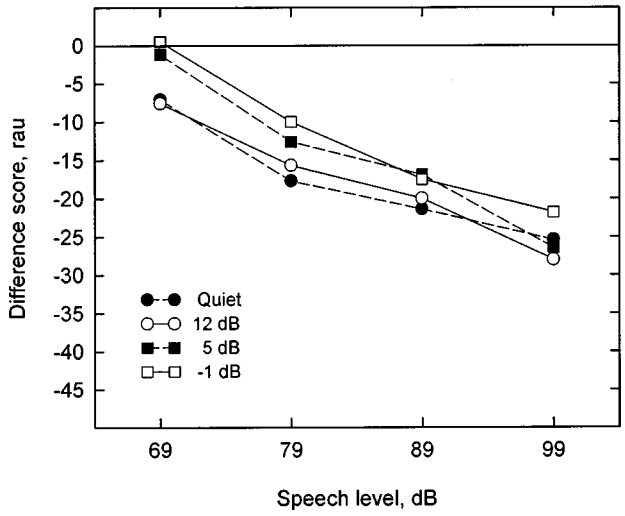


FIG. 7. Mean differences between observed and expected performance (in rau) for the hearing-impaired subjects after correction for differences in audibility over speech level and S/N ratio. The calculations assume the speech dynamic range equals 40 dB.

made no *a priori* assumptions about the dynamic range of the speech material.

The final steps were to convert the predicted proportionate scores into rau, to gain the advantages of the arcsine transform (Studebaker, 1985), and then to subtract these predicted values from the rau transforms of the mean obtained scores. The resulting difference scores, when plotted versus speech level, are PICA functions in that they represent the average effect of increasing speech and noise levels on speech-recognition performance after correction for the effects of audibility change.

C. PICA functions: Hearing impaired versus normal hearing

The PICA functions for the hearing-impaired subjects are reported in Fig. 7, while those for the normal-hearing subjects tested under the same conditions are shown in Fig. 8. In both cases, the results have been corrected for audibility effects. This operation had no effect on the slopes of the functions produced by the normal-hearing subjects because calculated audibility remained constant for these subjects. However, it did mitigate the effects of S/N ratio so that their PICA functions are clustered like those for the hearing-impaired subjects.

TABLE IV. Linear regression slopes (in rau/dB) for the PICA functions in Figs. 7 and 8.

S/N ratio	Normal hearing	Hearing impaired
Quiet	−0.22	−0.59
12 dB	−0.87	−0.66
5 dB	−0.96	−0.80
−1 dB	−0.73	−0.74

Slope values for the lines in Figs. 7 and 8 were calculated using linear regression and are reported in Table IV. These slopes, together with the figures, suggest that normal-hearing and hearing-impaired subjects produce similar results in noise when audibility changes are taken into account. In the case of the hearing-impaired subjects, it is also evident that the slope of the quiet function is only slightly less than the slopes of the noise functions. However, the quiet-function slope for the normal-hearing subjects differs substantially from the others.

For each figure, the overlap between the individual functions shows how well the prediction model accounts for the effects of S/N ratio on speech recognition. In the case of Fig. 7, the functions' proximity to the baseline at 69 dB reveal the model's ability to predict the hearing-impaired subjects' average absolute performance at the reference speech level. The combined results for both subject groups are shown in Fig. 9. In this case, the overlap of the results for the two subject groups reflects the model's ability to account for the average effects of hearing loss.

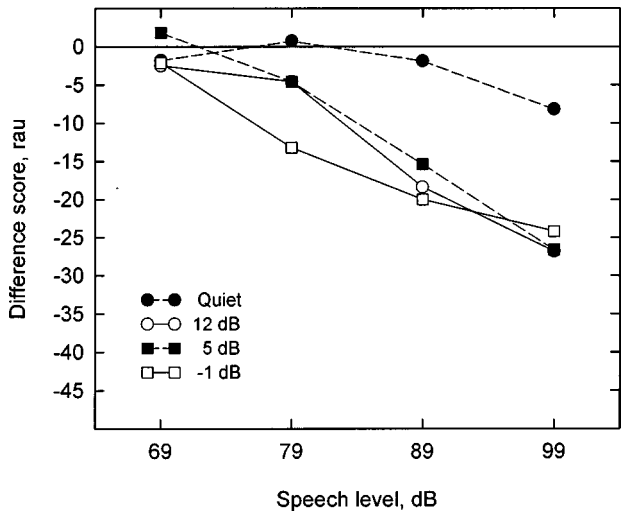


FIG. 8. Mean differences between observed and expected performance (in rau) for the normal-hearing subjects after correction for differences in audibility over speech level and S/N ratio. The calculations assume the speech dynamic range equals 40 dB.

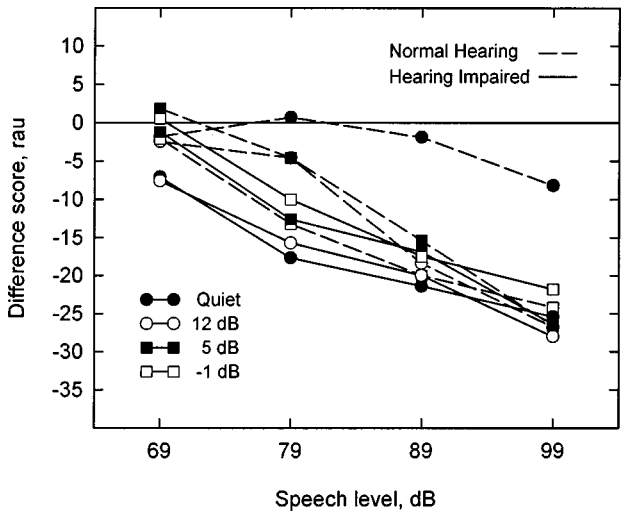


FIG. 9. Mean differences between observed and expected performance (in rau) for the normal-hearing subjects versus the hearing-impaired subjects. The data for both groups have been corrected for changes in audibility. The calculations assume the speech dynamic range equals 40 dB.

TABLE V. Intersubject standard deviations, in rau, associated with the unadjusted means for the 44 hearing-impaired subjects and 12 normal-hearing subjects tested under the same set of listening conditions. SPL refers to the long-term rms speech level.

SPL	S/N ratio			
	Quiet	12	5	−1
		Normal-hearing subjects		
69	5.5	8.4	7.0	8.9
79	8.3	8.9	8.7	6.3
89	6.7	11.9	8.1	13.4
99	7.0	10.7	10.7	13.4
		Hearing-impaired subjects		
69	26.9	24.1	21.0	15.3
79	23.7	19.0	16.5	15.4
89	18.6	15.3	16.2	12.7
99	17.5	14.3	14.4	15.1

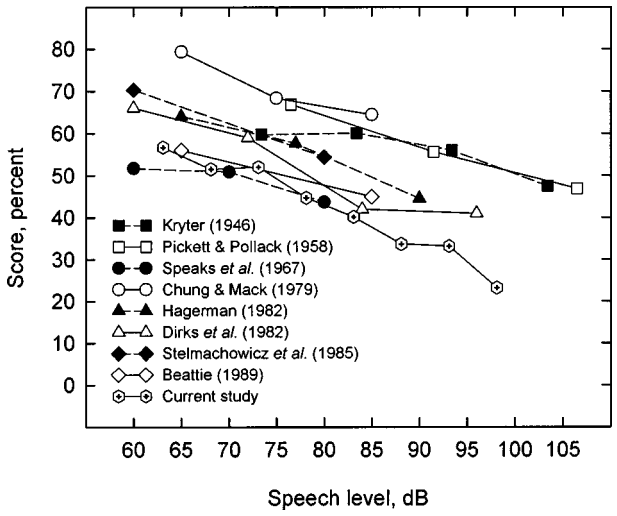


FIG. 10. PICA functions (in percent) from nine studies that tested normal-hearing subjects.

D. Intersubject variability

Table V reports intersubject standard deviations, in rau, for the 44 hearing-impaired subjects and the twelve normal-hearing subjects who were tested under the same conditions as the hearing-impaired subjects. The corresponding mean values are plotted in Figs. 2 and 6.

The values for the normal-hearing subjects at the lower speech and noise levels are about what one would expect for a 50-item word test (Thornton and Raffin, 1978; Studebaker, 1985; Studebaker *et al.*, 1995). There is a tendency for the intersubject differences to increase with both speech and noise levels. Some cell-to-cell inconsistency is seen in this pattern, probably because each of these values was based on just 12 scores.

In contrast to the normal-hearing subject trends, standard deviations for the hearing-impaired subjects are largest at the lowest speech and noise levels. This likely reflects the fact that these subjects had different audiometric configurations and that audiometric configuration has the largest effect on audibility when speech and noise levels are near threshold.

It is interesting that at the highest speech and noise levels, the intersubject variability of the hearing-impaired subjects is only modestly larger than that of the normal-hearing subjects. Such a result suggests that when the listening conditions are equivalent, the intersubject variability of normal-hearing and hearing-impaired subjects may be quite similar. However, this should not be considered conclusive because this experiment was not designed to answer questions about relative intersubject variability and does not lend itself to firm conclusions in this area.

III. DISCUSSION

A. Comparisons with other studies

The results of the present study indicate that even modestly high speech levels (>69 dB SPL) can produce substantial reductions in speech-recognition performance under conditions in which the S/N ratio remains constant. The biggest effects were observed at S/N ratios that produced scores near 50 rau, but reduced performance was also evident even when the S/N ratio exceeded 15 to 18 dB, the commonly presumed lower edge of the dynamic range of speech. These findings do not agree with those of earlier studies that have suggested there is a linear relationship between speech-threshold level and masker level through at least 90 dB SPL (e.g., Hawkins and Stevens, 1950; Hirsh and Bowman, 1953; Hirsh *et al.*, 1954; Duquesnoy and Plomp, 1983; Gordon-Salant, 1987). However, they do agree with the data from a number of other investigations.

Figure 10 shows the steepest PICA functions from this study and eight previous studies that used low-to-mid frequency emphasis maskers (Kryter, 1946; Pickett and Pollack, 1958; Speaks *et al.*, 1967; Chung and Mack, 1979; Dirks *et al.*, 1982; Hagerman, 1982; Stelmachowicz *et al.*, 1985; Beattie, 1989). All the data were obtained from normal-hearing subjects and have been plotted with speech level as the independent variable. The speech levels are in dB SPL and assume monaural listening under a supra-aural earphone. In those studies where the original data were not in this form (Kryter, 1946; Pickett and Pollack, 1958; Beattie, 1989, this study), the results were converted to SPL and/or moved along the *X*-axis by a correction factor estimated from data reported by Bentler and Pavlovic (1989) and ISO R226-1961 (ISO, 1961).[4]

The most important feature of Fig. 10 is the agreement among the slopes of the various functions. The different absolute performance levels only reflect the relative difficulty of the conditions that were used in each study. To provide single-number descriptors for these results, linear regression was used to calculate slopes (in %/dB) for each data set. These slopes are reported in Table VI, along with the type of speech and noise used in each study. Also included is a slope for one of the functions obtained by Pollack and Pickett (1958) in a study that used a white-noise masker.

Another way to evaluate data of this type is to calculate the change in S/N ratio required to maintain a constant level of performance. As noted earlier, Hawkins and Stevens (1950), Hirsh and Bowman (1953), Hirsh *et al.* (1954), and

TABLE VI. The steepest speech recognition PICA function slopes (%/dB) obtained in ten studies. Also reported are the speech and noise types used in each study and the data source (Fig./Table) and S/N ratio (dB) for each slope value.

Investigation	Speech	Noise	Source	S/N ratio	Slope
Kryter (1946)	PAL PB words	Engine spectrum	Fig. 3	0	−0.41
Pickett and Pollack (1958)	PAL PB words	−12 dB/oct	Fig. 2(B)	−10	−0.66
Pollack and Pickett (1958)	PAL PB words	White	Fig. 1	0	−0.48
Speaks *et al.* (1967)	Synthetic sentences	1-talker discourse	Fig. 2	−30	−0.40
Chung and Mack (1979)	W-22 words	Pink	Fig. 4(A)	12	−0.74
Dirks *et al.* (1982)	NU#6 words	12-talker babble	Fig. 1	0	−0.77
Hagerman (1982)	Swedish sentences	Talker spectrum	Fig. 6	varied	−0.79
Stelmachowicz *et al.* (1985)	NU#6 words	Broadband	Fig. 1	4	−0.80
Beattie (1989)	W-22 words	Auditec multitalker	Table II	6	−0.55
Current study	NU#6 words	Talker spectrum	Fig. 2	5	−0.92

Duquesnoy and Plomp (1983) all found that the S/N ratio needed to achieve a fixed performance level remained constant as a function of masker intensity level. In contrast, French and Steinberg (1947) reported an ''excess masking factor'' of 10 dB when the critical band sensation level (CBSL) reached 90 dB. According to French and Steinberg, their conclusions regarding this *m* factor were based partly on pure-tone masking data taken from a study by Fletcher and Munson (1937), and partly on their own data.

The rate at which masking increases in the French and Steinberg model ranges from about 1.15 dB of additional masking per dB of added noise (1.15 dB/dB) when the CBSL is just above 50 dB to about 1.40 dB/dB when the CBSL is near 90 dB. In 1958, Pickett and Pollack concluded that their results for CBSLs below 90 dB agreed with those of French and Steinberg (1947). However, they recommended a linear function for *m* that increased at a rate of 1.2 dB/dB whenever the CBSL exceeded 50 dB.

Table VII reports slope values for *m* based on data from 12 studies, including this one. In each case, the slope was calculated by plotting the change in speech versus noise level needed to maintain 50% performance. Inspection of the tabled values reveals that they vary markedly. Several studies observed little or no change in masking effectiveness with level; several found increases in the range of 1.1 to 1.2 dB/dB; and two reveal slopes of more than 1.3 dB/dB.

The cause of these discrepancies cannot be determined with absolute certainty, but it appears that larger effects occur whenever the masker spectrum matches the spectrum of the speech material and the speech type is either monosyllabic words or nonsense syllables. If Table VI is included in the evaluation, the data suggest that more consistent results are obtained when the dependent variable is the percentage score observed at a fixed S/N ratio as opposed to the change in S/N ratio needed to maintain a constant score. The relatively steep function found in the present study may reflect the combined influence of all three factors. Further research is needed to resolve these inconsistencies.

B. Data patterns

A comparison of the PICA functions for the various groups and conditions suggests that performance deteriorates at high speech and noise levels because of a relative increase in the effective masking level. However, this extra masking ''energy'' is not in the form of a ''fence'' located some number of dB below the level of the speech signal, as implied by the phrase ''equivalent noise of self masking'' used by French and Steinberg (1947). Nor is it simply an increase in the effective level of an existing applied noise based only on the level of that noise as in the case of the ''*m*'' factor of Fletcher and Munson (1937), French and Steinberg (1947),

TABLE VII. The change in S/N ratio needed to maintain a constant speech-recognition score, based on data from 12 studies. Also reported are the speech and noise types used in each study. The slope values indicate the amount of masking, in dB, per dB of applied noise.

Investigation	Speech	Noise	Slope (dB/dB)
Kryter (1946)	PAL PB words	Engine spectrum	1.12
French and Steinberg (1947)	Nonsense syllables	Quiet	1.15 to 1.4
Hawkins and Stevens (1950)	Continuous discourse	White	0
Hirsh and Bowman (1953)	Spondees	White	0
Hirsh *et al.* (1954)	Various	White	0
Pickett and Pollack (1958)	PAL PB words	−12 dB/oct	1.2
Dirks *et al.* (1982)	NU#6 words	12-voice babble	1.14 to 1.17
Duquesnoy and Plomp (1983)	Dutch sentences	Talker spectrum	0
Dubno *et al.* (1984)	SPIN-PI	12-voice babble	1.11
Stelmachowicz *et al.* (1985)	NU#6 words	Broadband	1.18
Walker and Byrne (1985)	Connected discourse	Talker spectrum	1.05
Current study	NU#6 words	Talker spectrum	1.34

and Pickett and Pollack (1958), or a reduction in the AI based only on the level of the speech signal as in the case of the ''*E*'' factor of Fletcher and Galt (1950).

Instead, there is evidence of an interaction between the speech and noise levels. When a speech signal is near normal conversational levels, and the noise level increases, performance falls at a rate that is closely predicted by the reduction in simple audibility. However, if the speech is at higher levels, and the noise level increases, then performance falls more rapidly. Similarly, when the noise is at a modest level, and the speech level increases, performance grows at a rate that is consistent with the improvement in simple audibility, while if the noise is at higher levels, and the speech level increases, then performance grows more slowly.

Further, when speech is presented in quiet, an increase in the level produces an apparent increase in the hypothetical internal noise. However, performance does not change until the internal noise level becomes high enough to reduce speech audibility. It follows that the negative effects of increasing speech level depend on the relationship between the applied or internal noise level and the bottom edge of the speech signal's dynamic range.

In the case of the normal-hearing subjects tested in the current study, their hypothesized internal noise levels fall well below the limits of the speech area assumed to contribute to intelligibility. When the speech was presented in quiet, therefore, increases in the speech level did not produce an effective masking level high enough to reduce performance until the speech was quite intense. On the other hand, when the speech was presented in noise, the noise level was close to or within the speech signal's dynamic range. Thus, an increase in the noise level raised the effective masking level, reduced speech audibility, and caused performance to decrease.

In the case of the hearing-impaired subjects, their hypothesized internal noise levels were higher and generally in the same intensity range as the levels of the applied noise. In addition, both types of noise were largely within the speech signal's dynamic range. Thus, any increase in masker effective level due to an increase in speech level produces a similar result whether in quiet or in applied noise.

C. Spread of masking

Outcomes like the ones in this study have often been attributed, at least partly, to ''spread of masking,'' (e.g., French and Steinberg, 1947; Danaher and Pickett, 1975; Rosenthal *et al.*, 1975; Milner *et al.*, 1984). We assessed this possibility by estimating the effect that the spread of masking of the noise masker should have had on the performance of the normal-hearing subjects tested in this study according to an algorithm published by Ludvigsen (1985). The same algorithm has been incorporated, without modification, into the speech intelligibility index (SII) standard (ANSI, 1997).

For cross-checking purposes, we used two entirely independent implementations of the Ludvigsen algorithm. One was a *C* program (Pavlovic, 1992) that contained the core elements of the SII. To evaluate the spread of masking component of that program by itself, it was necessary to turn off a speech-level correction called the ''Level distortion factor'' (LDF). The other method was to add the Ludvigsen algorithm to the LOTUS 1-2-3 spreadsheet that we had used earlier to make the predictions for this study. In both cases, Eq. (3) was used to convert the derived indexes into predicted NU#6 scores that could then be compared to the actual mean word scores.

TABLE VIII. Expected score reductions (in rau) from 69 to 99 dB SPL due to upward spread-of-masking (Masking alone) and spread-of-masking plus speech level distortion (Masking+LDF). The predicted values were derived using the calculation methods specified in the 1997 Speech Intelligibility Index (SII) standard (ANSI, 1997). The actual reductions produced by 12 normal-hearing subjects are also shown.

S/N	Masking alone	Masking+LDF	Subjects
Quiet	0.0	−12.9	−1.5
28	0.0	−12.9	−3.9
24	0.0	−13.0	−7.2
20	0.0	−13.2	−12.0
16	−1.6	−15.6	−17.6
12	−2.6	−17.5	−23.3
8	−3.7	−19.1	−27.5
5	−4.9	−20.0	−28.8
2	−6.0	−20.2	−28.4
−1	−7.8	−19.5	−26.2
−4	−9.1	−17.5	−22.6

When the same assumptions about the dynamic range were made in both programs, the predicted reductions in performance due to spread of the noise masking were within 1 rau of each other. Therefore, we decided to use only the results from the SII program in our comparisons with actual subject performance.

Table VIII reports the predicted versus the observed reduction in score that occurred between 64 and 99 dB SPL at each S/N ratio. Clearly, the actual changes were much larger than those predicted by the Ludvigsen algorithm. The calculated values are generally consistent with data from a number of previous studies that have examined the effects of spread of masking on speech recognition and that have shown those effects to be relatively small (Gordon-Salant, 1984; Milner *et al.*, 1984; Dorman *et al.*, 1985; Goshorn and Studebaker, 1994).

The third column of Table VIII shows the predicted effects of higher-than-normal speech and noise levels when the LDF is included in the calculations. Notice that the reductions in performance are much larger than those for spread of masking alone. This indicates that the authors of S3.5-1997 thought speech level ''distortion,'' rather than spread of masking, was the main cause of lower scores at high levels. However, the predicted effects of level in quiet are also bigger than the subject data revealed. Thus, our data suggest an interaction between speech level and noise level, whereas the SII LDF assumes only speech level is important.

D. Hearing status

Another issue considered by this study was whether high presentation levels have the same effects on hearing-impaired subjects that they do on normal-hearing subjects. At first glance, based on the uncorrected mean data, it would seem they do not. The two groups clearly did not perform the same; also, high levels appear to have had a greater negative

impact on the performance of the normal-hearing subjects. However, these differences in group performance are, to a very large extent, simply the result of the different ways in which audibility changed with level for the two groups. This is revealed by the fact that when the data were corrected for such differences, the results for the hearing-impaired subjects were similar to the results for the normal hearers tested in noise.

Of course, the accuracy of this interpretation is based at least partly on the validity of the method used to correct the hearing-impaired data for audibility loss. It was for this reason that we used a relatively simple calculation method in which the only factor obtained on a *post hoc* basis was the assumed dynamic range. And, as stated previously, this factor did not substantially affect those aspects of the data (the PICA function slopes) that provide the major evidence for our conclusions.

In addition, the proposed interpretation is consistent with published findings from the literature. For example, studies by Humes *et al.* (1987), Zurek and Delhorne (1987), Dubno and Schaefer (1992), and Ching *et al.* (1998) all report that the effects of level, under comparable conditions of audibility, are the same for normal-hearing and hearing-impaired subjects. The study by Ching *et al.* (1998) is of particular interest because it evaluated a wide range of hearing losses and used wholly different prediction and test methods than this study; yet it still found that ''...the negative effect of high signal level on speech performance was similar for normal-hearing and hearing-impaired listeners.'' (p. 1134).

E. Age

The results of this study suggest that the negative effects of increasing speech and noise levels on speech recognition are similar for all adult subjects, up to at least 80 years of age. On average, the hearing-impaired subjects below age 70 and those 70 years and over performed virtually the same when tested in noise. In quiet, a small difference between the speech-recognition scores of the two age groups was noted, but this could very easily have been due to differences in the hearing-threshold levels of the two subject groups. Moreover, after correction for audibility effects, the results for both older groups were very similar to those for the much younger normal-hearing subjects.

F. Dynamic range

The audibility corrections used in this study assume that the dynamic range of speech is 40 dB. This value is considerably larger than the 30-dB range that has been widely used in the past; therefore, it deserves additional discussion. However, before beginning, it should be noted that the dynamic range assumed in these calculations has only a small effect on the main factor that underlies most of our conclusions about the data—that is, the *slopes* of the PICA functions. For example, for the normal-hearing subjects, the PICA function slopes did not change except when we used a dynamic range greater than 43 dB to predict performance in quiet; then, the slope changed only slightly. The slopes for the hearing-impaired subjects were more affected by the size of the range, but not by very much. For the quiet data, the slope change was only 0.025 rau/dB for each dB of change in the assumed dynamic range; for the noise data, it was even smaller. Thus, it seems unlikely that our conclusions, based on slopes, were significantly altered by the use of a larger dynamic range.

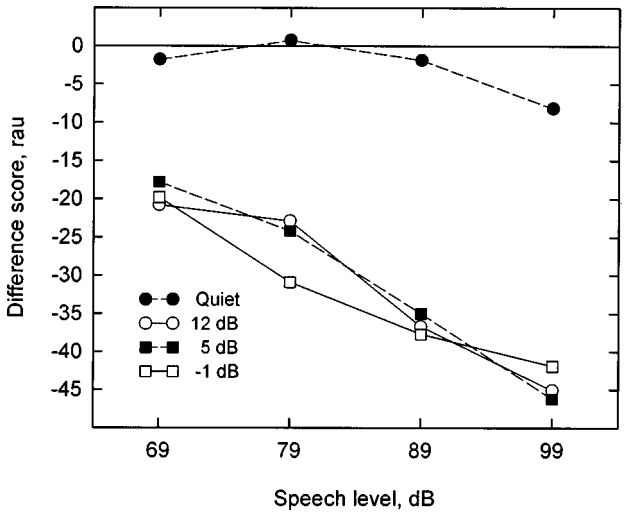


FIG. 11. Mean differences between observed and expected performance (in rau) for the normal-hearing subjects after correction for differences in audibility over speech level and S/N ratio. The calculations assume the speech dynamic range equals 30 dB.

Nevertheless, the dynamic range assumed in AI-type calculations is an important matter that has gone largely unchallenged for many years. The present study provides several pieces of evidence concerning this issue. One is illustrated by Fig. 11. The results plotted in this figure are based on the same normal-hearing subject data as those in Fig. 8. The only difference is that the results in Fig. 8 assume a dynamic range of 40 dB, while those in Fig. 11 assume the conventional 30 dB.

Figure 11 clearly reveals the problem that occurs when a 30-dB dynamic range is used to predict the performance of normal hearers. This is that a particular set of transfer function coefficients and other constants provide good predictions in quiet, or in noise, but not under both conditions. For example, in Fig. 11, the difference between the observed and predicted scores for the quiet condition is generally less than 5 rau, but the masked results are overpredicted by 20 rau or more.

As it is unacceptable to use different transfer functions and/or other constants for different listening conditions, this problem appeared at first glance to be unsolvable. However, calculations revealed that the disagreement between the quiet and masked results was substantially diminished when dynamic ranges exceeding 30 dB were used in the prediction scheme. In the present case, increasing the dynamic range to 40 dB had virtually no effect on the predictions for the quiet data, but substantially improved the predictions for the masked results (compare Figs. 8 and 11). At the same time, it improved predictions of the average hearing-impaired subject data as well.

In addition to this outcome, at least two other findings

from this study indicate that the dynamic range is more than 30 dB. One is that the normal-hearing subjects who were tested in quiet produced slightly higher scores when the speech level increased from 64 to 79 dB SPL. This suggests that audibility was still increasing under conditions where the conventional assumptions indicated it was not. The other pertinent result is that the normal-hearing subjects performed noticeably poorer in low-level noise (16 to 28 dB S/N) than they did in quiet and the difference between these conditions increased with speech level. Theoretically, this should not have occurred if the dynamic range is only 15 to 18 dB below the rms level of the speech, as assumed in ANSI, 1997 and ANSI, 1969, respectively, because the noise under these assumptions is below the limits of the range and therefore unable to reduce speech audibility. In contrast, if the range is at least 28 dB below the speech signal's rms level, as suggested by the present data, and the rms peaks of speech exceed the long-term rms level of speech by 12 to 15 dB, then the dynamic range must be at least 40 to 43 dB.

Credible evidence for a larger dynamic range can also be found in the literature. For example, the AI procedures of French and Steinberg (1947) used a dynamic range of 36 dB, while those of Fletcher and Galt (1950) used a dynamic range of 68 dB.[5] More recently, De Gennaro *et al.* (1981) reported data indicating the physical dynamic range of speech is as much as 50 dB. Levitt (1982) has also pointed out that the results of Dunn and White (1940), which form the principal basis for the assumption of a 30-dB dynamic range, were limited by equipment noise and that the actual dynamic range of speech may be closer to 40 or 50 dB. Finally, results from a study of the intensity importance function by Studebaker *et al.* (1997) are consistent with a range as large as the one used in this study. Taken together, these various facts suggest that the true dynamic range of speech may be considerably larger than the value now in common use. Clearly, further study of this important issue is needed.

IV. CONCLUSIONS

Based on the results of this study, and others in the literature, the following conclusions were reached.

(1) Speech intelligibility in noise decreases when speech levels exceed 69 dB SPL and the S/N ratio remains constant.
(2) The effects of speech and noise level are synergistic. That is, the negative effects of added noise level are greater when the speech level is high, and vice versa.
(3) The deterioration in intelligibility can be modeled as a relative increase in the effective masking level including that produced by the hypothetical internal noise that determines threshold in quiet.
(4) Normal-hearing and hearing-impaired subjects are affected similarly by increased speech and noise levels when differences in speech audibility are considered.
(5) The negative effects of increasing speech and noise levels on speech recognition are similar for all adult subjects, at least up to 80 years of age.
(6) The effective dynamic range of speech may be larger than the commonly assumed value of 30 dB.

ACKNOWLEDGMENTS

This project was supported by research grant No. 5 R01 DC 00154-16 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health, and by funds from the Center for Research Initiatives and Strategies for the Communicatively Impaired (CRISCI). The authors want to thank Christine G. Eubanks and Ginger A. Gray for their help with subject recruitment and data collection. Two anonymous reviewers also offered many valuable suggestions.

[1]The level of the hypothesized internal noise in any given band is calculated by adding the SPL of the pure-tone threshold for the band-center frequency to the bandwidth, in dB, and then subtracting the corresponding normal critical ratio.

[2]The +16 dB S/N ratio condition was presented at both test sites as a control measure. The data from all the subjects were included in the average for this condition. Because only mean values were compared, and only compared descriptively, the larger *N* at this S/N did not bias the results.

[3]Other curvilinear functions also might have been used to describe the distribution of PICA function slopes. We chose the Gaussian function because it fit the data and because a Gaussian-like distribution of slope values is an expected outcome whenever the range of possible performance values is restricted on both ends of the distribution of scores.

[4]The speech levels reported by Kryter (1946) were increased by 8.4 dB. This value equals the difference between the average of the minimum audible pressure coupler (MAPC) values and the average of the 6 cc to *Z* Coupler values at 250–4000 Hz (Bentler and Pavlovic, 1989). The speech levels reported by Pollack and Pickett (1958) were increased by 1.5 dB. This value equals the difference between the average of the minimum audible field (MAF) monaural values (Bentler and Pavlovic, 1989) and the average of the MAF binaural values (ISO, 1961) at 250–4000 Hz. The speech levels reported by Beattie (1989) were converted from HL to SPL by adding 20 dB. The speech levels in this study were decreased by 0.9 dB. This value equals the average difference between the 6 cc and *Z* coupler values at 250–4000 Hz (Bentler and Pavlovic, 1989).

[5]Comparisons with the ranges used by French and Steinberg (1947) and Fletcher and Galt (1950) are not straightforward. Both used functions that varied with level. Also, Fletcher and Galt used two functions: one that applied to the signal as a whole and another (with the same range) that applied to the level of each band compared to a specially defined average speech level.

ANSI (**1969**). ANSI S3.5-1969, ''American National Standards Methods for the Calculation of the Articulation Index'' (American National Standards Institute, New York).

ANSI (**1989**). ANSI S3.6-1989, ''Specifications for Audiometers'' (American National Standards Institute, New York).

ANSI (**1997**). ANSI S3.5-1997, ''American National Standards Methods for the Calculation of the Speech Intelligibility Index'' (American National Standards Institute, New York).

Beattie, R. C. (**1989**). ''Word recognition functions for the CID W-22 test in multitalker noise for normally hearing and hearing-impaired subjects,'' J. Speech Hear. Disord. **54**, 20–32.

Bentler, R. A., and Pavlovic, C. V. (**1989**). ''Transfer functions and correction factors used in hearing aid evaluation and research,'' Ear Hear. **10**, 58–63.

Ching, T. Y. C., Dillon, H., and Byrne, D. (**1998**). ''Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification,'' J. Acoust. Soc. Am. **103**, 1128–1140.

Chung, D. Y., and Mack, B. (**1979**). ''The effect of masking by noise on word discrimination scores in listeners with normal hearing and with noise-induced hearing loss,'' Scand. Audiol. **8**, 139–143.

Danaher, E. M., and Pickett, J. M. (**1975**). ''Some masking effects produced by low frequency vowel formants in persons with sensorineural hearing loss,'' J. Speech Hear. Res. **18**, 261–271.

De Gennaro, S., Braida, L. D., and Durlach, N. I. (**1981**). ''A statistical analysis of third-octave speech amplitude distributions,'' paper presented at the 101st meeting of the Acoustical Society of America, Ottawa, Ontario, Canada.
Dirks, D. D., Kamm, C. A., Dubno, J. R., and Velde, T. M. (**1981**). ''Speech recognition performance at loudness discomfort level,'' Scand. Audiol. **10**, 239–246.
Dirks, D. D., Morgan, D. E., and Dubno, J. R. (**1982**). ''A procedure for quantifying the effects of noise on speech recognition,'' J. Speech Hear. Disord. **47**, 114–123.
Dorman, M. F., Lindholm, J. M., and Hannley, M. T. (**1985**). ''Influence of the first formant on the recognition of voiced stop consonants by hearing-impaired listeners,'' J. Speech Hear. Res. **28**, 377–380.
Dubno, J. R., Dirks, D. D., and Morgan, D. E. (**1984**). ''Effects of age and mild hearing loss on speech recognition in noise,'' J. Acoust. Soc. Am. **76**, 87–96.
Dubno, J. R., and Schaefer, A. B. (**1992**). ''Comparison of frequency selectivity and consonant recognition among hearing-impaired and masked normal-hearing listeners,'' J. Acoust. Soc. Am. **91**, 2110–2121.
Dunn, H. K., and White, S. D. (**1940**). ''Statistical measurements on conversational speech,'' J. Acoust. Soc. Am. **11**, 278–288.
Duquesnoy, A. J., and Plomp, R. (**1983**). ''The effect of a hearing aid on the speech-reception threshold of hearing-impaired listeners in quiet and noise,'' J. Acoust. Soc. Am. **73**, 2166–2173.
Fabry, D. A., and Van Tasell, D. J. (**1986**). ''Masked and filtered simulation of hearing loss: effects on consonant recognition,'' J. Speech Hear. Res. **29**, 170–178.
Fletcher, H. (**1922**). ''The nature of speech and its interpretation,'' Bell Syst. Tech. J. **1**, 129–144.
Fletcher, H., and Munson, W. A. (**1937**). ''Relation between loudness and masking,'' J. Acoust. Soc. Am. **9**, 1–10.
Fletcher, H., and Galt, R. H. (**1950**). ''The perception of speech and its relation to telephony,'' J. Acoust. Soc. Am. **22**, 89–151.
French, N. R., and Steinberg, J. C. (**1947**). ''Factors governing the intelligibility of speech sounds,'' J. Acoust. Soc. Am. **19**, 90–119.
Gordon-Salant, S. (**1984**). ''Effects of reducing low-frequency amplification on consonant perception in quiet and noise,'' J. Speech Hear. Res. **27**, 483–493.
Gordon-Salant, S. (**1987**). ''Age related differences in speech recognition performance as a function of test format and paradigm,'' Ear Hear. **8**, 277–281.
Goshorn, E. L., and Studebaker, G. A. (**1994**). ''Effects of intensity on speech recognition in high- and low-frequency bands,'' Ear Hear. **15**, 454–460.
Hagerman, B. (**1982**). ''Sentences for testing speech intelligibility in noise,'' Scand. Audiol. **11**, 79–87.
Hargus, S. E., and Gordon-Salant, S. (**1995**). ''Accuracy of speech intelligibility index predictions for noise-masked young listeners with normal hearing and for elderly listeners with hearing impairment,'' J. Speech Hear. Res. **38**, 234–243.
Hawkins, J. E., and Stevens, S. S. (**1950**). ''The masking of pure tones and of speech by white noise,'' J. Acoust. Soc. Am. **22**, 6–13.
Hirsh, I. J., and Bowman, W. D. (**1953**). ''Masking of speech by bands of noise,'' J. Acoust. Soc. Am. **25**, 1175–1180.
Hirsh, I. J., Reynolds, E. G., and Joseph, M. (**1954**). ''The intelligibility of different speech materials,'' J. Acoust. Soc. Am. **26**, 530–538.
Humes, L. E., Dirks, D. D., Bell, T. S., and Kincaid, G. E. (**1987**). ''Recognition of nonsense syllables by hearing-impaired listeners and by noise-masked normal hearers,'' J. Acoust. Soc. Am. **81**, 765–773.
ISO (**1961**). ISO/R226-1961, ''Normal equal-loudness contours for pure tones and normal threshold of hearing under free field listening conditions'' (American National Standards Institute, New York).
Jerger, J., and Jerger, S. (**1971**). ''Diagnostic significance of PB word functions,'' Arch. Otolaryngol. **93**, 573–580.
Kryter, K. D. (**1946**). ''Effects of ear protective devices on the intelligibility of speech in noise,'' J. Acoust. Soc. Am. **18**, 413–417.
Levitt, H. (**1982**). ''Speech discrimination ability in the hearing impaired: spectrum considerations,'' in *The Vanderbilt Hearing Aid Report*, edited by G. A. Studebaker and F. H. Bess (Monographs in Contemporary Audiology, Upper Darby, PA), pp. 32–43.
Ludvigsen, C. (**1985**). ''Relations among some psychoacoustic parameters in normal and cochlearly impaired listeners,'' J. Acoust. Soc. Am. **78**, 1271–1280.
Matesich, J. S. (**1991**). Monosyllabic word test presentation program. Unpublished computer program (The University of Memphis, Memphis, TN).
Matesich, J. S. (**1993**). Monosyllabic word test scoring program. Unpublished computer program (The University of Memphis, Memphis, TN).
Milner, P., Braida, L. D., Durlach, N. I., and Levitt, H. (**1984**). ''Perception of filtered speech by hearing-impaired listeners,'' Appendix to: Braida, L. D. (**1984**). ''Articulation testing methods for evaluating speech reception by impaired listeners,'' in ASHA Reports No. 14, *Speech Recognition by the Hearing Impaired*, edited by E. Elkins (American Speech-Language-Hearing Association, Rockville, MD).
Pavlovic, C. V. (**1992**). SII software, V3.0, 17 Dec. 1992.
Pickett, J. M., and Pollack, I. (**1958**). ''Prediction of speech intelligibility at high noise levels,'' J. Acoust. Soc. Am. **30**, 955–963.
Pollack, I., and Pickett, J. M. (**1958**). ''Masking of speech by noise at high sound levels,'' J. Acoust. Soc. Am. **30**, 127–130.
Rosenthal, R. D., Lang, J. K., and Levitt, H. (**1975**). ''Speech reception with low-frequency speech energy,'' J. Acoust. Soc. Am. **57**, 949–955.
Speaks, C., Karmen, J. L., and Benitez, L. (**1967**). ''Effect of a competing message on synthetic sentence identification,'' J. Speech Hear. Res. **10**, 390–396.
Steeneken, H. J. M., and Houtgast, T. (**1980**). ''A physical method for measuring speech-transmission quality,'' J. Acoust. Soc. Am. **67**, 318–326.
Stelmachowicz, P. G., Jesteadt, W., Gorga, M. P., and Mott, J. (**1985**). ''Speech perception ability and psychophysical tuning curves in hearing-impaired listeners,'' J. Acoust. Soc. Am. **77**, 620–627.
Studebaker, G. A. (**1985**). ''A rationalized arcsine transform,'' J. Speech Hear. Res. **28**, 455–462.
Studebaker, G. A., McDaniel, D. M., and Sherbecoe, R. L. (**1995**). ''Evaluating relative speech recognition performance using the proficiency factor and rationalized arcsine differences,'' J. Am. Acad. Audiol. **6**, 173–182.
Studebaker, G. A., and Sherbecoe, R. L. (**1991**). ''Frequency-importance and transfer functions for recorded CID W-22 word lists,'' J. Speech Hear. Res. **34**, 427–438.
Studebaker, G. A., Sherbecoe, R. L., and Gilmore, C. (**1993**). ''Frequency-importance and transfer functions for the Auditec of St. Louis recordings of the NU-6 word test,'' J. Speech Hear. Res. **36**, 799–807.
Studebaker, G. A., Sherbecoe, R. L., and Gwaltney, C. A. (**1997**). ''The development of a monosyllabic word intensity importance function,'' Poster paper presented at the Second Biennial NIDCD/VA Hearing Aid Research and Development Conference.
Thornton, A. R., and Raffin, M. J. M. (**1978**). ''Speech-discrimination scores modeled as a binomial variable,'' J. Speech Hear. Res. **21**, 507–518.
Walker, G., and Byrne, D. (**1985**). ''Reliability of speech intelligibility estimation for measuring speech reception thresholds in quiet and noise,'' Aust. J. Audiol. **7**, 23–31.
Zurek, P. M., and Delhorne, L. A. (**1987**). ''Consonant reception in noise by listeners with mild and moderate sensorineural hearing impairment,'' J. Acoust. Soc. Am. **82**, 1548–1559.

Forward masking among infant and adult listeners

Lynne A. Werner
Department of Speech and Hearing Sciences, University of Washington, 1417 N.E. 42nd Street, Seattle, Washington 98105-6246

(Received 27 March 1998; revised 21 July 1998; accepted 17 December 1998)

Psychophysical forward-masked thresholds were estimated for 3- and 6-month-old infants and for adults. Listeners detected a repeated 1000-Hz probe, with 16-ms rise time, no steady-state duration, and 16-ms fall time. Unmasked thresholds were determined for one group of listeners who were trained to respond when they heard the probe but not at other times. In the masking conditions, each tone burst was preceded by a 100-ms broadband noise masker at 65 dB SPL. Listeners were trained to respond when they heard the probe and masker, but not when they heard the masker alone. The masker–probe interval, Δt, was either 5, 10, 25, or 200 ms. Four groups of subjects listened in the masked conditions, each at one value of Δt. Each listener attempted to complete a block of 32 trials including four probe levels chosen to span the range of expected thresholds. ''Group'' thresholds, based on average psychometric functions, as well as thresholds for individual listeners, were estimated. Both group and individual thresholds declined with Δt, as expected, for both infants and adults. Infants' masked thresholds were higher than those of adults, and comparison of masked to unmasked thresholds suggested that infants demonstrate more forward masking than adults, particularly at short Δt. Forward masking appeared to have greater effects on 3-month-olds' detection than on either 6-month-olds' or adults'. Compared to adults, 6-month-olds demonstrated more forward masking only for Δt of 5 ms. Thus, susceptibility to forward masking may be nearly mature by 6 months of age. © *1999 Acoustical Society of America.* [S0001-4966(99)00704-3]

PACS numbers: 43.66.Dc [JWH]

INTRODUCTION

The purpose of the experiment reported here was to examine the development of recovery from adaptation, as measured by thresholds under forward masking. Adaptation is one of several mechanisms believed to limit temporal resolution, the ability to follow amplitude changes in a sound over time. Immaturity of susceptibility to adaptation has been proposed as an explanation of age-related changes in temporal resolution.

Several studies indicate that temporal resolution undergoes development during infancy and childhood. Irwin *et al.* (1985) first reported that children's ability to detect temporal interruptions, or gaps, in a low-frequency noise band did not mature until 10 years of age. Only slight improvements in gap detection in high-frequency noise bands were observed between 6 years and adulthood. Wightman *et al.* (1989) subsequently reported that gap detection was immature among 3- and 4-year-old children, but mature by age 5 years for noise bands centered at both 400 and 2000 Hz. Werner *et al.* (1992) found that gap detection was very poor among 3-, 6-, and 12-month-old infants. Studies of duration discrimination among infants and 4–10-year-olds (Elfenbein, Small, and Davis, 1993; Jensen and Neff, 1993; Morrongiello and Trehub, 1987) reported similar age-related improvements in performance. By 6 years of age, many children in these studies could discriminate between sounds of different durations as well as adults could.

Although age-related change in measures of temporal resolution has been documented, the mechanisms responsible for that change are unknown. A listener's ability to follow a sound's amplitude envelope is actually limited by several underlying mechanisms. In order to accurately encode the amplitude envelope, the time of occurrence of amplitude peaks must be accurately represented in the pattern of neural discharges. This ability is the one most closely identified with temporal resolution. Encoding the amplitude envelope, however, also requires precise intensity resolution: The neural response must change sufficiently with a change in sound amplitude for the difference in amplitude to be detectable. Furthermore, the neural representation of the amplitude envelope will be affected by adaptation, the process that is responsible for the rapid reduction in primary neural response rate that occurs just after sound onset, and for a reduction in sensitivity that occurs following offset of a sound. Finally, even frequency resolution may theoretically be involved in temporal resolution insofar as narrow filters ''ring'' for a longer time period than broad filters do. Immaturity in any of these underlying mechanisms could contribute to immature perception of temporal modulations of sound.

The development of three of the four mechanisms just described has been examined. There is growing evidence that temporal coding is mature before gap-detection thresholds (Hall and Grose, 1994; Levi and Werner, 1995, 1996), although whether temporal coding is immature in early infancy is unsettled (Levi, Folsom, and Dobie, 1993, 1995; Levi and Werner, 1996). Intensity coding, however, may develop over a time course similar to that seen for gap detection: Intensity discrimination is immature until 5 or 6 years of age (Jensen and Neff, 1993; Maxon and Hochberg, 1982; Sinnott and Aslin, 1985). Frequency resolution is only immature at high frequencies at 3 months of age and is mature by 6 months of

age (Hall and Grose, 1991; Olsho, 1985; Schneider, Morrongiello, and Trehub, 1990; Spetner and Olsho, 1990).[1]

Trehub, Schneider, and Henderson (1995) have suggested recently that infants and children are better at detecting gaps in very short duration sounds than they are at detecting gaps in very long duration sounds, because they are more susceptible to adaptation effects than adults. Evoked potential rate-effect studies suggest, in fact, that the time required to recover from prior stimulation is longer for infants and children than it is for adults, and the time required to recover from prior stimulation has been shown to be directly related to susceptibility to adaptation in single-unit studies (e.g., Harris and Dallos, 1979). Increasing the rate of stimulus presentation has a greater effect on the auditory brainstem response (ABR) among infants younger than 3 months of age than it does on adults' ABR (Dey-Sigman, Ruth, and Rubel, 1984; Fujikawa and Weber, 1977; Lasky, 1984), and immature, pronounced rate effects for middle- and long-latency evoked potentials are reported well into childhood (Hall, 1992; Jerger *et al.*, 1987; Kraus *et al.*, 1985). Along the same lines, Lasky (1991, 1993) reported that newborns' click-evoked ABRs are more affected by a forward masker than are those of adults.

Forward masking is thought of as a psychophysical measure of adaptation in the auditory system. The relationship between masker–probe delay and the probe's threshold is an indication of the time course of the auditory system's recovery from prior stimulation. The precise locus and mechanism of forward masking are not completely understood (e.g., Shannon, 1990; Turner, Relkin, and Doucet, 1994), and under certain circumstances masker–probe confusions can elevate forward-masked thresholds (Neff, 1985). If, however, infants are more susceptible to adaptation than adults, we would expect that they would demonstrate more pronounced forward masking effects than adults.

I. METHOD

The probe was a 1000-Hz tone, with 16-ms rise and fall times and no steady-state duration. The masker was a broadband noise, 100 ms in duration, including 16-ms rise and fall times. Masker level was set at 65 dB SPL. A broadband noise masker was used to make sure that the masker and probe were distinct in sound quality, so that masker–probe confusions would be less likely (Neff, 1985). The masker–probe interval, Δt, was 5, 10, 25, or 200 ms, measured from the offset of the masker to the onset of the probe. The stimuli were presented to each listener's right ear through an Etymotic ER-1 insert earphone in a foam ear tip, trimmed as necessary to fit infant ear canals. Testing was carried out in a double-walled booth. The experiment was controlled by a computer.

The subjects were 114 3-month-old infants; 58 6-month-old infants; and 51 18–30-year-old adults. Infants were tested within 2 weeks of their 3-month or 6-month birthday. Adult age was taken as age at last birthday. All subjects passed a screening questionnaire assessing risk for hearing loss (Joint Committee on Infant Hearing, 1991), had neither been treated for otitis media within 2 weeks nor had more than two episodes of otitis media, and passed screening tympanometry at each laboratory visit. None had previously participated in a psychoacoustic study. An additional 49 3-month-olds and 21 6-month-olds were tested, but did not provide data because they responded inconsistently to the probe.

The number of infants who were tested but did not provide data is troublesome. It is not typical of published studies using similar methods, and no explanation for the high attrition rate is apparent. These data were collected after 2 years of pilot testing to optimize the details of the procedure; no procedural variant was successful in producing data from more infants than the one used in the final study. Four different individuals tested infants for this study over a period of 3 years. None was more successful than the others; the attrition rate did not decline with time or practice. The data obtained from adults compare as expected to those reported for well-trained listeners, arguing that a peculiarity of the stimuli was not responsible. No differences between infants who provided data and those who did not—in age or gender, for example—have been identified.

Despite the difficulties in interpretation resulting from the apparently select sample, I believe that the data presented here provide a meaningful representation of infants' auditory capacities for several reasons. First, these data are indistinguishable, in terms of variability and internal consistency, from data obtained in other tasks (e.g., detection in quiet, frequency discrimination, masked detection) for which a similar number of trials was obtained from each infant. Second, the infants who provided data in this study took no longer to meet training criteria than did infants who have provided data for other tasks. Third, comparison of the data of infants who provided complete data sets to those who provided partial data sets indicated little difference in level of performance.

Of course, it is impossible to know what the results would have been had the data of a higher proportion of infant subjects been included. It is unlikely, however, that the performance of the infants who were excluded would have been better than that of the infants who were included. Thus, it would be reasonable to think of the present results as establishing a lower limit on infants' threshold under forward masking; that is, it would be safe to say that infants are at least as bad at detecting tones under forward masking as indicated by the present results.

Initially, data collection concentrated on 3-month-olds and adults, and ten thresholds/condition were obtained at each age. When it became apparent that there was a large difference between 3-month-olds and adults, 6-month-olds were tested to see whether any age-related change in forward masking occurred during infancy. In the interest of timely dissemination, and because it was difficult to obtain these data, it was decided to obtain five thresholds/condition from 6-month-olds. Changes in the data analysis procedure after data collection was completed resulted in changes in the number of subjects contributing data at different ages and in different conditions.

Unmasked detection data were provided by 45 3-month-olds; 31 6-month-olds; and 10 adults. The data from infants were originally described by Werner and Marean (1991).

They were collected over the same time period as the masked-detection data, using the same methods, and were reanalyzed for comparison with the masked-detection data here. Some of these subjects provided as many as 60 test trials; all data obtained from a given subject were used in the present analysis. For masked detection, each subject listened at one value of Δt. The numbers of 3-month-olds providing data were 14, 16, 18, and 21 at Δt of 5, 10, 25, and 200 ms, respectively. The corresponding numbers of 6-month-olds were 5, 5, 7, and 10. For adults, the corresponding numbers were 11, 9, 8, and 13. Because of the small number of subjects tested at 6 months, the results for that age group must be interpreted cautiously.

Listeners heard two types of trials during the experiment, signal and no-signal trials. In the unmasked-detection condition, signal trials consisted of 20 repetitions of the probe with a 444-ms interstimulus interval. Unmasked no-signal trials were periods of equal duration during which no stimulus was presented. In masked-detection conditions, masker and probe were presented on signal trials, while only the masker was presented on no-signal trials. On signal trials, 20 repetitions of the sequence masker–Δt–probe were presented, with 344 ms between repetitions. On no-signal trials, 20 repetitions of the noise burst were presented, with interstimulus intervals of $372\,\text{ms}+\Delta t$. The interval between noise-burst onsets and the total trial duration were the same for masked signal and masked no-signal trials in each Δt condition. These long trials were used because 3-month-olds often have long response latencies.

The method used to assess infants' sensitivity was an observer-based procedure (Werner, 1995). The infant was seated on a parent's lap in the test booth, facing a window and a video camera. An assistant sat to the infant's left and manipulated quiet toys to keep the infant attending at midline. The parent and assistant listened to masking sounds to ensure that they could not hear any sound that was presented to the infant. Two mechanical toys with lights in a dark Plexiglass box were placed to the infant's right. An observer watched the infant though the window and on a video monitor. When the infant was quiet and attentive, the observer cued the computer to begin a trial. The observer did not know whether a signal or no-signal trial was being presented. Signal and no-signal trials were presented with equal probability. The observer decided, based on the infant's behavior, whether a signal or no-signal trial was being presented. If the observer correctly identified a signal trial before the trial ended, one of the mechanical toys in the booth was activated and illuminated to reinforce the infant's response. The observer received feedback after every trial. Infant responses typically include head or eye movements or changes in activity. Success in this procedure clearly depends on the infant's response to the probe and on the observer's ability to detect the infant's response, but if the observer can reliably identify signal trials, then the infant must be detecting the probe.

The infant/observer team was required to meet a criterion level of performance prior to data collection. Initially, the level of the probe was fixed at a value expected to be clearly audible to the listener, and the mechanical toy reinforcer was activated at the end of every signal trial. The purpose of this procedure was to demonstrate the association between the probe and the reinforcer to the infant. Once the infant/observer team had achieved four of five consecutive trials correct (hits or correct rejections), another series of trials was completed in which the probe was still clearly audible, but the reinforcer was only activated if the observer correctly identified a signal trial. This phase of the experiment continued until the infant/observer team had achieved four of the last five no-signal trials correct and four of the last five signal trials correct. The average number of trials required to complete the training procedure was 22.7 (s.d. $=8.1$) for 3-month-olds and 23.3 trials (s.d.$=9.3$) for 6-month-olds, with no differences across conditions.

The test phase of the experiment consisted of a block of 32 trials. On a signal trial, one of four probe levels was presented. The order of probe levels was randomized with the constraint that equal numbers of trials were completed at each level by the end of the 32 trials. The probe levels were evenly spaced over a 20-dB range and were chosen for each age group on the basis of pilot testing. In the 200-ms Δt condition, some infants were tested with a range of levels centered around the expected unmasked threshold, while others were tested with a range of levels centered 10 dB higher. The additional condition was added to ensure that the range of levels used did not strongly influence the threshold obtained.

Adults listened alone in the booth. They were told to respond whenever they heard ''the sound that makes the toy come on.'' Otherwise, the procedures were identical to those used with the infants. Adults nearly always completed the training procedure in the minimum possible number of trials and completed all test trials.

If an infant completed at least 30 test trials, an individual threshold estimate was attempted. At the time the data were collected, the plan was to exclude infants with false alarm rates greater than 0.25 or with nonmonotonic psychometric functions. It was subsequently decided to deal with high false alarm rate and nonmonotonic psychometric functions statistically rather than by exclusion. However, some infants had been retested and provided more than one block of test trials, because the false alarm rate was high or because performance was nonmonotonic with level in the first block completed. In such cases, all of the infant's data were used to estimate a single threshold. Psychometric functions of a form described by Green (1995) were fit to the data using a maximum likelihood criterion. The underlying relationship between level and the proportion of ''yes'' responses, $p(\text{''yes''})$, was assumed to be linear, but bounded by the false alarm rate and an upper ''asymptote.'' Given the small number of trials per subject, it was unlikely that the data would be better described by a more complex function. In fact, logistic and probit functions resulted in very similar threshold estimates, but far fewer successful fits than a linear function. The false alarm rate was estimated for each subject on the basis of responses on the 16 no-signal trials. The upper bound was taken as 0.86 for infants and as 0.97 for adults, values obtained by Bargones, Werner, and Marean (1995) using the same probe. Threshold was taken as the

level at which $p(\text{``yes''})=0.6$, the approximate midpoint of the infant psychometric function.

Thirty-nine unmasked thresholds and 58 masked thresholds (11–19 per condition) were obtained from 3-month-olds. Twenty-eight unmasked and 23 masked thresholds (3–11 per condition) were obtained from 6-month-olds. Ten unmasked and 49 masked thresholds (8–13 per condition) were obtained from adults. The subjects who provided data, but not thresholds, were evenly divided between those who completed fewer than 30 test trials and those who completed 30 test trials but for whom a threshold could not be estimated (psychometric function with negative or zero slope, fitting procedure would not converge).

A bias-free measure of sensitivity would be clearly preferable to $p(\text{``yes''})$, particularly because infants tend to have higher false alarm rates than adults do. However, in previous work we have found that the use of the parametric statistic d' may not be justified in observer-based studies of infant psychoacoustic performance (Werner, Kopyar, and Bargones, 1993), and it is unfortunate that no means for estimating a nonparametric bias-free sensitivity statistic were available in the laboratory at the time that these data were collected. Furthermore, when performance in this study was described in terms of d', the number of individual psychometric functions that could be successfully fit was reduced by 50%–70%. For those reasons, the decision was made to present the results in terms of $p(\text{``yes''})$. However, analysis in terms of $d'$ yielded essentially the same results as those obtained using $p(\text{``yes''})$, and to provide the proper perspective, equivalent values of d' are reported in the text.

II. RESULTS

Because the number of trials obtained even from infants who completed the test procedure was small, the data were analyzed in two ways. First, ''group'' psychometric functions were constructed combining the data from all subjects of a given age in a given condition. Group psychometric functions have frequently been used in developmental studies where few data have been obtained from each subject (e.g., Trehub *et al.*, 1995). Group psychometric functions have the advantage of allowing us to use the data of any subject who completed even one test trial. Group psychometric functions have the limitation that they provide no information about variability in sensitivity across subjects, so individual thresholds were also estimated for as many infants as possible. Individual thresholds were only estimated for infants who completed at least 30 test trials. While the individual thresholds are based on few trials, they can be interpreted in light of the group thresholds, and they provide some information about the variability in sensitivity across subjects.

A. Group psychometric functions and thresholds

The data of all subjects who completed any test trials were used to construct group psychometric functions for each listening condition (masked detection at four values of Δt and unmasked detection). About 14% of the infants included completed fewer than 32 test trials; all of the adult subjects had complete data sets. For 3-month-olds, the percentages of trials included from incomplete data sets were 0, 5, 12, 8, and 8% for the masking conditions with Δt of 5, 10, 25, and 200 ms and the unmasked condition, respectively. The corresponding percentages for the 6-month-olds were 29, 15, 11, 7, and 8%. The proportion of ''yes'' responses recorded at each probe level and on no-signal trials across all subjects who completed test trials was calculated. These functions are shown in Fig. 1.

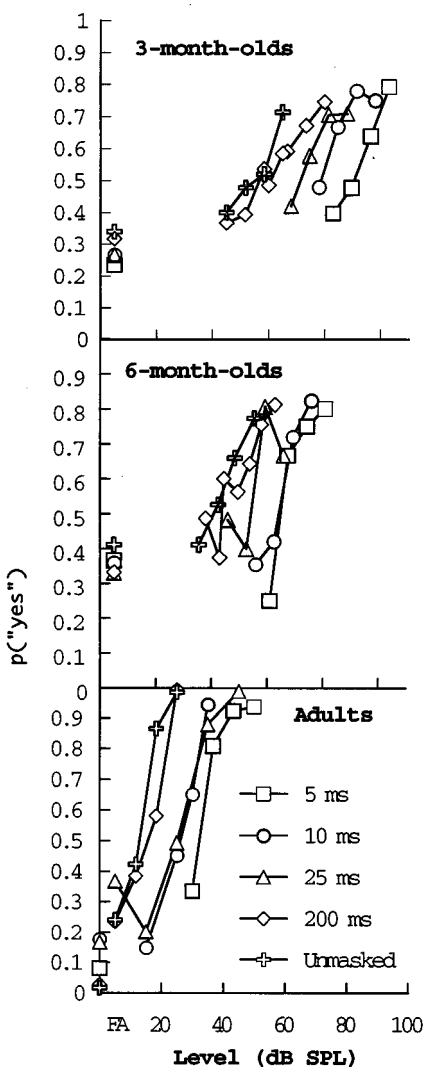


FIG. 1. Group psychometric functions for detection of unmasked and forward masked 1000-Hz tone pips, for three age groups. Points plotted over ''FA'' are false alarm rates.

The proportion of ''yes'' responses increased with probe level more or less monotonically at all three ages and in all listening conditions. The functions of 6-month-olds had more nonmonotonic points than those of 3-month-olds or adults, but recall that there were fewer 6-month-old infants tested. Among 3-month-olds average $p(\text{``yes''})$ on no-signal trials, false alarms, ranged from 0.24–0.34. The false alarm rate of 6-month-olds was a little higher, 0.33–0.41. The adults' false alarm rates were 0.02–0.18 in masked detection and 0.01 in unmasked detection. These rates are typical of those generally observed for infants and adults in this procedure. On average, $p(\text{``yes''})$ at the highest level tested in

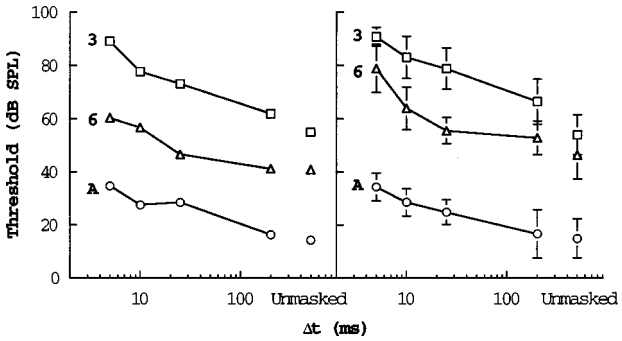


FIG. 2. Forward-masked threshold as a function of Δt for three age groups, calculated by two methods. Average thresholds estimated from the group psychometric functions are plotted in the left panel. Thresholds calculated for individual listeners are averaged and plotted in the right-hand panel. ''3'' indicates the data of 3-month-olds; ''6'' indicates the data of 6-month-olds; and ''A'' indicates the data of adults. Error bars indicate ± 1 sem. Unmasked thresholds are the points plotted to the right in each panel.

each condition was around 0.8 for the infants and 1.0 for the adults. That would mean that infants achieved an average d' of 1.0–1.6 at the highest level tested in each condition, while adults achieved an average d' between 2.5 and 3.9. Finally, notice that performance was poor among infants who were tested at lower levels in forward-masked detection at 200-ms Δt. These are the lowest four data points on the infants' functions for 200 ms. Thus, the shift in the group psychometric function to higher levels for masked detection in this condition is not simply a result of using a higher range of probe levels.

The group psychometric functions tend to shift to higher levels for masked detection than for unmasked detection, as expected. The functions also shift, more or less, to higher levels in masked detection as Δt decreases. From the spread of the psychometric functions, it appears that 3-month-olds' detection is affected by the forward masker more than adults' detection is.

The thresholds calculated from the group average psychometric functions will be referred to as ''group thresholds.'' To examine the age difference in forward masking quantitatively, a group threshold was calculated for each psychometric function. Threshold was taken as the level at which p(''yes'') was equal to 0.6, the approximate midpoint of the psychometric functions. This point was estimated by fitting a line by least-squares criterion to each psychometric function. The group thresholds obtained are plotted in Fig. 2 (left panel). In general, the 3-month-olds have higher group thresholds than the 6-month-olds who have higher group thresholds than the adults under all conditions.[2] The infant–adult difference is much larger than typically observed for long-duration tone detection. That result is consistent with previous observations from this laboratory (Werner and Marean, 1991; Bargones *et al.*, 1995), but is not understood. Berg and her colleagues (e.g., Berg, 1991, 1993; Berg and Boswell, 1995) have reported the same differences between 6-month-olds and adults in thresholds for short-duration sounds under several conditions. Because the focus of the current paper is on the effect of a forward masker on threshold, rather than on the absolute values of the thresholds, the reader interested in possible reasons for the large age differences in threshold is referred to those published papers.

It also appears that the difference between 3-month-olds' and adults' group thresholds is greater under forward masking than it is when the probe is unmasked; in fact, the difference between 3-month-olds' and adults' group thresholds increases progressively with decreasing Δt. These trends are not evident in the group-threshold difference between 6-month-olds and adults. The effect of introducing a forward masker and the decline of forward masking with increasing Δt seem quite similar for 6-month-olds and adults. At Δt of 200 ms, forward-masked threshold is about the same as unmasked threshold for 6-month-olds and for adults. For 3-month-olds, the masked-group threshold at 200 ms is about 5 dB higher than unmasked-group threshold. The same pattern is evident in thresholds calculated from group psychometric functions based on d' using a threshold criterion of $d' = 0.51$, equivalent to a hit rate of 0.6 and a false alarm rate of 0.4. The only difference was that the d' thresholds of the 6-month-olds tend to be about 6 dB closer to those of the 3-month-olds.

That masking appears to grow at a faster rate with decreasing Δt in the youngest listeners suggests that 3-month-olds are more susceptible to forward masking than adults are. Adults and 6-month-olds appear to be similar in their susceptibility to forward masking.

B. Individual thresholds

Psychometric functions were fit to the data of individual infants who completed at least 30 test trials, as described above. The average individual unmasked thresholds and masked thresholds as a function of Δt are shown in Fig. 2 (right panel). Most of the expected trends are evident in the average individual thresholds as they were in the group thresholds. Masked threshold declines with increasing Δt at all ages. Both masked and unmasked thresholds decline progressively with age. The average individual thresholds of 3-month-olds and of adults are very similar to their respective group thresholds in all conditions. The average individual thresholds of 6-month-olds are consistently higher than group thresholds for that age. Average individual unmasked thresholds and masked thresholds for Δt from 10 to 200 ms were 5–10 dB higher than the corresponding group thresholds, but the average individual masked threshold at a Δt of 5 ms was nearly 20 dB higher than the corresponding group threshold among 6-month-olds. The difference appears to be due not to an artifact of the threshold calculation, but to between-infant differences in sensitivity. The difference between 3-month-olds' and adults' masked thresholds is greater than the difference between their unmasked thresholds, and the difference between their masked thresholds tends to increase with decreasing Δt. The threshold differences between 6-month-olds and adults is larger for a Δt of 5 ms than other conditions, but the difference is similar for Δt of 10–200 ms and the unmasked condition. Individual thresholds calculated from d' showed parallel results, with the exceptions that no thresholds could be calculated for 6-month-olds at Δt of 5 ms and that the thresholds of the 6-month-olds were somewhat closer to those of the 3-month-olds.

An analysis of variance (ANOVA) of threshold, with age and listening condition (four masking conditions and one unmasked condition) as factors, confirms these impressions. The age-by-listening condition interaction, the main effect of age, and the main effect of listening conditions were all highly significant (all p's<0.001). Within-age-group analyses of the effect of listening condition were used to explore the nature of the age by listening condition interaction. Among 3-month-olds, the effect of listening condition was significant ($p<0.001$), and all pairwise comparisons between listening conditions were significant (p's<0.001) except for the 10 vs 25 ms comparison ($p>0.05$). In other words, masked thresholds were higher than unmasked threshold for all Δt, and all but one increase in Δt lead to a significant decrease in threshold. Among 6-month-olds, the effect of listening condition was also significant ($p<0.001$). Masked threshold was significantly higher than unmasked threshold at all Δt. Masked thresholds at 200 and 25 ms and at 25 and 10 ms did not differ. Masked threshold at 5 ms was significantly higher than all other thresholds. Among adults, masked threshold at 200 ms was not different from unmasked threshold. All other masked thresholds were significantly higher than unmasked and the 200-ms masked threshold. While the 25- and the 5-ms masked thresholds were significantly different, the 10- and 25-ms and the 5- and 10-ms masked thresholds were not. Thus, introducing a forward masker at a Δt of 200 ms leads to an increase in threshold among infants, but not among adults. More significant increases in masked threshold are observed as Δt is decreased among 3-month-olds than among 6-month-olds or adults. This is consistent with a greater growth of forward masking with decreasing Δt among 3-month-olds than among older listeners.

C. Amount of masking

A more direct way to examine infants' relative susceptibility to forward masking would be to compare the amount of masking (masked minus unmasked threshold) produced by a forward masker across age groups. Amount of masking could not be calculated for the listeners tested here, because masked and unmasked thresholds were not obtained from the same listeners. However, an approximation to the amount of masking could be obtained by subtracting the average unmasked threshold for each age group from the average masked threshold of that age group.

The average amount of masking calculated from the group thresholds and from individual thresholds is plotted in Fig. 3. The estimates derived from group thresholds are within 5 dB of those derived from individual thresholds, with the exception of the 6-month-olds at 5 ms, for whom the individual threshold estimate indicates considerably more masking than does the group threshold estimate. Three-month-olds show more masking than either 6-month-olds or adults at all values of Δt by both estimates. On the other hand, 6-month-old and adult amounts of masking are very similar in all cases by the group threshold estimates and very similar at all Δt except 5 ms by the individual threshold estimates. The individual threshold estimate of 6-month-olds' amount of masking at Δt of 5 ms is higher than that estimate for adults.

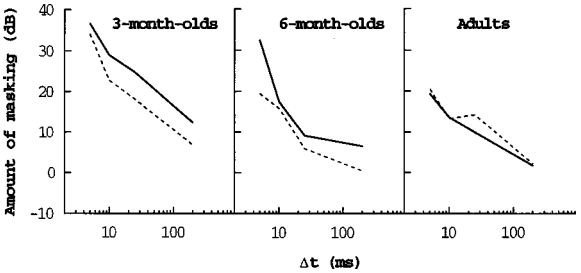


FIG. 3. Average amount of masking estimated by subtracting average unmasked threshold from average masked threshold at each Δt, for three age groups. The dashed lines indicate values calculated from thresholds estimated from the group psychometric functions. The solid lines indicate values calculated from average individual thresholds.

One can calculate ''amount of masking'' for each subject using the individual thresholds by subtracting the mean unmasked threshold for the appropriate age group from each masked threshold. An age X Δt ANOVA with pairwise *post hoc* comparisons of these values showed that 3-month-olds have significantly higher amounts of masking than adults and 6-month-olds at all values of Δt, while the 6-month-olds only have significantly higher amounts of masking than adults at Δt of 5 ms. The problem with this analysis is that the variance in the amount of masking estimated will simply equal the variance in masked threshold. The variance in ''real'' amounts of masking, calculated from masked and unmasked thresholds from the same individuals, would be expected to be higher than that of the masked (or unmasked) thresholds. Thus, a greater number of significant differences might be found for the estimates of amount of masking based on average unmasked threshold than for the individual amounts of masking. The significant differences between 3-month-olds and older listeners in estimated amounts of masking, then, may overstate the extent of the difference. However, following the same logic, one would still not expect 6-month-olds' individual amounts of masking to differ from those of adults.

III. DISCUSSION

The results of this experiment indicate that at least until 3 months of age, infants are more susceptible to forward masking than are adults, particularly at shorter values of Δt. Three-month-olds also appear to be susceptible to forward masking at longer masker–probe intervals than adults are. Susceptibility to forward masking appears to decline, however, between 3 and 6 months of age. By 6 months of age, it may be mature. All of these conclusions, of course, must be tempered by the limitations of the methods used to derive them. At the same time, the data obtained here are consistent with those in previous reports.

The forward-masking data obtained from the adult comparison group in this study are roughly comparable to those reported previously (e.g., Elliott, 1962a, 1962b; Jesteadt, Bacon, and Lehman, 1982). Probably the closest match in terms of stimulus conditions are the data of Elliott (1962b), who measured forward masking of a 10-ms, 1000-Hz probe by a

100-ms, broadband noise masker at 70 dB SPL, for Δt ranging from 0 to 50 ms. Elliott reported about 12 dB of masking at 5 ms, declining to about 4 dB at 25 ms. The adult listeners here were estimated to exhibit about 19 dB of masking at 5 ms, declining to about 8–10 dB at 25 ms. Thus, the current sample of adults exhibits somewhat more masking than Elliott's subjects, a result not surprising given their limited experience with the task.

Because a fixed masker level was used for all subjects and because infants' absolute thresholds are higher than those of adults, it is reasonable to consider how differences in sensation level may have contributed to the age differences observed in forward masking. Reducing the sensation level of the masker has two effects: a reduction in the amount of masking and a reduction in the slope of the function relating amount of masking to masker–probe interval (e.g., Jesteadt *et al.*, 1982). Thus, the amount of masking shown in Fig. 3 may underestimate infants' susceptibility to forward masking. Whether or not this is the case depends on the locus of the immaturity in absolute sensitivity relative to the locus of the immaturity in susceptibility to forward masking. For example, if the primary cause of the immaturity in absolute sensitivity is conductive, then the effective level of the forward masker would certainly be lower for the infants, and their relative susceptibility to forward masking would be greater than indicated by the amount of masking in Fig. 3. If immaturity in absolute sensitivity arises rostral to the site(s) which limit detection under forward masking, then the level of the forward masker would be effectively the same at all ages. Given that neither the mechanisms underlying forward masking nor the site of immaturities in absolute sensitivity have been established, it is impossible to draw a conclusion about this issue. However, assuming that the inner ear is mature, peripheral conductive immaturity would be expected to attenuate the sound in the region of 1000 Hz by about 3 dB for 3- or 6-month-olds (Keefe *et al.*, 1993). That difference in effective masker level would be expected to produce a very small difference in the amount of masking produced. Moreover, the fact that infants' function relating amount of masking to masker–probe interval is either steeper than or the same as that of adults argues against sensation level as a primary explanation for the observed age differences in forward masking.

The 3-month-olds' poor performance under forward-masking conditions is striking. There is a substantial difference between 3-month-olds and adults in threshold and in amount of forward masking, no matter which infant data one chooses for the comparison. Lasky's (1991) report is the only previous study of forward masking in human infants. Lasky reported that an 82-dB SPL broadband noise with Δt of 10 ms elevated click-evoked ABR thresholds of 1–3-day-old infants 6 dB more than it elevated adult thresholds. This is a considerably smaller age difference than the 10–20-dB difference observed psychophysically with a less intense masker and older infants. Thus, it is unlikely that immaturity at or peripheral to the auditory brainstem can account for much of the psychophysical age difference. It is still possible that immaturity of neural structures central to the brainstem is involved. Unfortunately, the existing literature on the development of evoked potentials originating in the auditory thalamus or cortex (e.g., Jerger *et al.*, 1987; Kraus *et al.*, 1985; Stapells *et al.*, 1988), while consistent with the idea that the infant's neural response is more susceptible to the effects of prior stimulation, does not permit direct comparison to either Lasky's or the present results. Jerger *et al.* (1987) reported that among 2–6-month-old infants, a slow positive peak in the middle latency response (MLR) was only observed when 500-Hz tone bursts were presented at rates less than 4/s. Adult MLRs are routinely recorded at rates greater than 10/s (e.g., Kraus *et al.*, 1985), but the difference in sensitivity to rate of stimulation is not readily converted to an age difference in amount of forward masking. In any case, mounting evidence that forward masking reflects retrocochlear processes in adults (Shannon, 1990; Turner *et al.*, 1994), and recent demonstrations of the relationship between evoked potential and perceptual development (Werner, Folsom, and Mancl, 1993, 1994) buttress the hypothesis that primary neural immaturities beyond the level of the brainstem may underlie immaturity in forward masking.

On the other hand, it seems clear that 6-month-old infants are less susceptible to forward masking than are 3-month-olds. In fact, 6-month-olds appear to be nearly mature in their susceptibility to forward masking, with the possible exception of forward masking at very short Δt. Given the high variability in 6-month-olds' performance at 5-ms Δt and the small number of 6-month-olds providing data, it is difficult to interpret the results in this condition. With reference to the evoked potential studies just cited, however, it should be noted that 6-month-olds may well show adultlike forward masking effects, although their auditory evoked potentials remain immature.

That infants exhibit more forward masking than adults is consistent with the hypothesis of Trehub *et al.* (1995) that infants are more susceptible to adaptation effects than adults are. Trehub *et al.* used this hypothesis to explain why infants appear to be so much more immature in gap detection when they are tested using continuous stimuli rather than very short duration tone pips. One aspect of the present results that is not consistent with that idea is the fact that while infant susceptibility to forward masking decreases between 3 and 6 months of age, infant gap-detection thresholds do not change over the same age period (Werner *et al.*, 1992).

As is always the case in the study of development, the possibility that immaturity of processes outside the primary auditory pathway influences infants' performance in a psychophysical test must be considered. Several authors have argued that age differences in thresholds of the order of magnitude reported here cannot in any case be accounted for by general inattention (e.g., Schneider and Trehub, 1992; Wightman and Allen, 1992; Viemeister and Schlauch, 1992). It is clear in any case that a general tendency to inattention cannot account for the greater susceptibility to forward masking seen in infants, for if infants were simply off-task a certain proportion of the time, we might expect that their unmasked and masked thresholds would be similarly affected (see also Bargones *et al.*, 1995). That infants took about as many trials to learn to respond to the probe in the

presence of a forward masker as they did to the unmasked probe argues against the idea that infants tended to be more inattentive in the masking conditions. Moreover, it is not clear why inattentiveness would be greater for shorter Δt. Similarly, if the difference between 3- and 6-month-olds were due to methodological limitations, then those limitations would have to differentially affect masked over unmasked thresholds.

Werner and Bargones (1991) showed that the presence of a spectrally distant, simultaneous noise raised tone-detection thresholds among 6-month-old infants, but not among adults, by 5–10 dB. This type of masking is often referred to as informational masking. If infants are unable to ignore a temporally distant masker, just as they are unable to ignore a simultaneous masker that is spectrally distant from the probe, then they might appear to be especially susceptible to the effects of a forward masker. In other words, failure to limit listening to the expected time of occurrence of the probe would add noise to the decision process and increase masked threshold. Neff and her colleagues have clearly shown that adults' thresholds for a tone can be elevated by the simultaneous presentation of multiple, spectrally distant, random-frequency tones (Neff, 1995; Neff and Callaghan, 1988). Allen and Wightman (1995) have recently reported that thresholds are unmeasurable for most 3–4-year-olds when a single, spectrally distant, random-frequency tone is presented simultaneously with a probe, although the effect of the random-frequency tone is similar for adults and children from whom thresholds can be obtained. Werner and Bargones' result suggests that infants have difficulty ignoring a simultaneous masker, even when its frequency is certain. In the present context, it is interesting that adults can readily ignore multiple, spectrally distant, random-frequency forward maskers (Neff, 1991). It may be that 3-month-olds cannot ignore a forward masker in any condition. It appears that 6-month-olds can efficiently limit listening to the appropriate time, as they, like adults, exhibit little or no forward masking when the probe follows the masker by at least 200 ms.

It is difficult to choose between these hypotheses on the basis of the current data. For example, if adding a forward masker, at any Δt, raises infants' thresholds above unmasked by some amount, but the function relating Δt to masked threshold for infants simply parallels that of adults, then an informational masking account may be more tenable than a primary auditory system explanation. In the current data, unfortunately, the group threshold analysis seems to suggest a more dramatic decline in forward masking with increasing Δt among 3-month-olds, but not 6-month-olds, while the individual threshold analysis shows 6-month-olds, but not 3-month-olds, with more dramatic declines in forward masking relative to adults. We may not be able to obtain sufficient psychophysical data from infants to discern such details. However, it may be possible to use other measures, such as evoked potentials, to try to determine the source of early immaturities in forward masking.

ACKNOWLEDGMENTS

This research was supported by grant No. R01 DC00396 from the National Institute of Deafness and Other Communication Disorders, National Institutes of Health. The data described in this paper were collected by Janelle Constantino, Lisa Mancl, Cam Marean, and Julianne Siebens.

[1]In any case, were frequency resolution to be poorer among infants or children than among adults, better temporal resolution would be predicted for young listeners.

[2]Psychometric functions including only complete data sets were compared to those including all test data to examine the possibility that incomplete data sets biased the group threshold estimates. For 3-month-olds, there was essentially no effect of including incomplete data sets. For 6-month-olds, group thresholds were 2–4 dB lower in three forward-masked conditions ($\Delta t = 5$, 10, and 200 ms) when the incomplete data sets were excluded, but those in the other two conditions were not affected. Thus, the inclusion of incomplete data sets appeared to have little impact on group threshold estimates.

Allen, P., and Wightman, F. (**1995**). ''Effects of signal and masker uncertainty on children's detection,'' J. Speech Hear. Res. **38**, 503–511.

Bargones, J. Y., Werner, L. A., and Marean, G. C. (**1995**). ''Infant psychometric functions for detection: Mechanisms of immature sensitivity,'' J. Acoust. Soc. Am. **98**, 99–111.

Berg, K., and Boswell, A. (**1995**). ''Temporal summation of 500-Hz tones and octave-band noise bursts in infants and adults,'' Percept. Psychophys. **57**, 183–190.

Berg, K. M. (**1991**). ''Auditory temporal summation in infants and adults: Effects of stimulus bandwidth and masking noise,'' Percept. Psychophys. **50**, 314–320.

Berg, K. M. (**1993**). ''A comparison of thresholds for 1/3-octave filtered clicks and noise bursts in infants and adults,'' Percept. Psychophys. **54**, 365–369.

Dey-Sigman, S. E., Ruth, R. A., and Rubel, E. W. (**1984**). ''ABR rate/intensity interaction: Developmental effects.'' Paper presented at the annual meeting of the American Speech-Hearing-Language Association, Washington, D.C., November, 1984.

Elfenbein, J. L., Small, A. M., and Davis, M. (**1993**). ''Developmental patterns of duration discrimination,'' J. Speech Hear. Res. **36**, 842–849.

Elliott, L. L. (**1962a**). ''Backward and forward masking of probe tones of different frequencies,'' J. Acoust. Soc. Am. **34**, 1116–1117.

Elliott, L. L. (**1962b**). ''Backward masking: Monotic and dichotic conditions,'' J. Acoust. Soc. Am. **34**, 1108–1115.

Fujikawa, S. M., and Weber, B. A. (**1977**). ''Effects of increased stimulus rate on brainstem electric response (BER) audiometry as a function of age,'' J. Am. Aud. Soc. **3**, 147–150.

Green, D. M. (**1995**). ''Maximum likelihood procedures and the inattentive listener.'' J. Acoust. Soc. Am. **97**, 3749–3760.

Hall, III, J. W., and Grose, J. H. (**1991**). ''Notched-noise measures of frequency selectivity in adults and children using fixed-masker-level and fixed-signal-level presentation,'' J. Speech Hear. Res. **34**, 651–660.

Hall, III, J. W. (**1992**). *Handbook of Auditory Evoked Responses* (Allyn and Bacon, Boston).

Hall, III, J. W., and Grose, J. H. (**1994**). ''Development of temporal resolution in children as measured by the temporal modulation transfer function,'' J. Acoust. Soc. Am. **96**, 150–154.

Harris, D. M., and Dallos, P. (**1979**). ''Forward masking of auditory nerve fiber responses,'' J. Neurophysiol. **42**, 1083–1107.

Irwin, R. J., Ball, A. K. R., Kay, N., Stillman, J. A., and Rosser, J. (**1985**). ''The development of auditory temporal acuity in children,'' Child Dev. **56**, 614–620.

Jensen, J. K., and Neff, D. L. (**1993**). ''Development of basic auditory discrimination in preschool children,'' Psychol. Sci. **4**, 104–107.

Jerger, J., Chmiel, R., Glaze, D., and Frost, J. D. J. (**1987**). ''Rate and filter dependence of the middle-latency response in infants,'' Audiology **26**, 269–283.

Jesteadt, W., Bacon, S. P., and Lehman, J. R. (**1982**). ''Forward masking as a function of frequency, masker level, and signal delay,'' J. Acoust. Soc. Am. **71**, 950–962.

Joint Committee on Infant Hearing (**1991**). ''1990 Position statement,'' ASHA **33**, Suppl. 5, 3–5.

Keefe, D. H., Bulen, J. C., Arehart, K. H., and Burns, E. M. (**1993**). ''Ear-canal impedance and reflection coefficient in human infants and adults,'' J. Acoust. Soc. Am. **94**, 2617–2638.
Kraus, N., Smith, D., Reed, N. L., Stein, L. K., and Cartee, C. (**1985**). ''Auditory middle latency responses in children: effects of age and diagnostic category,'' Electroencephalogr. Clin. Neurophysiol. **62**, 343–351.
Lasky, R. E. (**1984**). ''A developmental study on the effect of stimulus rate on the auditory evoked brain-stem response,'' Electroencephalogr. Clin. Neurophysiol. **59**, 411–419.
Lasky, R. E. (**1991**). ''The effects of rate and forward masking on human adult and newborn auditory evoked response thresholds,'' Dev. Psychobiol **24**, 21–64.
Lasky, R. E. (**1993**). ''The effect of forward masker duration, rise/fall time, and integrated pressure on auditory brain stem evoked responses in human newborns and adults,'' Ear Hear. **14**, 95–103.
Levi, E., and Werner, L. A. (**1995**). ''Modulation detection of sinusoidally amplitude-modulated (SAM) noise in 3- and 6-month-old infants: Preliminary data,'' Abstracts Midwinter Research Meeting ARO **18**, 56.
Levi, E. C., Folsom, R. C., and Dobie, R. A. (**1993**). ''Amplitude-modulation following response (AMFR): Effects of modulation rate, carrier frequency, age, and state,'' Hearing Res. **68**, 42–52.
Levi, E. C., Folsom, R. C., and Dobie, R. A. (**1995**). ''Coherence analysis of envelope-following responses (EFRs) and frequency-following responses (FFRs) in infants and adults,'' Hearing Res. **89**, 21–27.
Levi, E. C., and Werner, L. A. (**1996**). ''Amplitude modulation detection in infancy: Update on 3-month-olds,'' Abstracts Midwinter Research Meeting ARO **19**, 142.
Maxon, A. B., and Hochberg, I. (**1982**). ''Development of psychoacoustic behavior: Sensitivity and discrimination,'' Ear Hear. **3**, 301–308.
Morrongiello, B. A., and Trehub, S. E. (**1987**). ''Age related changes in auditory temporal perception,'' J. Exp. Child Psychol. **44**, 413–426.
Neff, D. L. (**1985**). ''Stimulus parameters governing confusion effects in forward masking,'' J. Acoust. Soc. Am. **78**, 1966–1976.
Neff, D. L. (**1991**). ''Forward masking by maskers of uncertain frequency content,'' J. Acoust. Soc. Am. **89**, 1314–1323.
Neff, D. L. (**1995**). ''Signal properties that reduce masking by simultaneous, random-frequency maskers,'' J. Acoust. Soc. Am. **98**, 1909–1920.
Neff, D. L., and Callaghan, B. P. (**1988**). ''Effective properties of multicomponent simultaneous maskers under conditions of uncertainty,'' J. Acoust. Soc. Am. **83**, 1833–1838.
Olsho, L. W. (**1985**). ''Infant auditory perception: Tonal masking,'' Inf. Beh. Dev. **7**, 27–35.
Schneider, B. A., Morrongiello, B. A., and Trehub, S. E. (**1990**). ''The size of the critical band in infants, children, and adults,'' J. Exp. Psychol. [Hum. Percept.] **16**, 642–652.
Schneider, B. A., and Trehub, S. E. (**1992**). ''Sources of developmental change in auditory sensitivity,'' in *Developmental Psychoacoustics*, edited by L. A. Werner and E. W. Rubel (American Psychological Association, Washington, DC), pp. 3–46.
Shannon, R. V. (**1990**). ''Forward masking in patients with cochlear implants,'' J. Acoust. Soc. Am. **88**, 741–744.
Sinnott, J. M., and Aslin, R. N. (**1985**). ''Frequency and intensity discrimination in human infants and adults,'' J. Acoust. Soc. Am. **78**, 1986–1992.
Spetner, N. B., and Olsho, L. W. (**1990**). ''Auditory frequency resolution in human infancy,'' Child Dev. **61**, 632–652.
Stapells, D. R., Galambos, R., Costello, J. A., and Makeig, S. (**1988**). ''Inconsistency of auditory middle latency and steady-state responses in infants,'' Electroencephalogr. Clin. Neurophysiol. **71**, 289–295.
Trehub, S. E., Schneider, B. A., and Henderson, J. (**1995**). ''Gap detection in infants, children, and adults,'' J. Acoust. Soc. Am. **98**, 2532–2541.
Turner, C. W., Relkin, E. M., and Doucet, J. (**1994**). ''Psychophysical and physiological forward masking studies: probe duration and rise-time effects,'' J. Acoust. Soc. Am. **96**, 795–800.
Viemeister, N. F., and Schlauch, R. S. (**1992**). ''Issues in infant psychoacoustics,'' in *Developmental Psychoacoustics*, edited by L. A. Werner and E. W. Rubel (American Psychological Association, Washington, DC), pp. 191–210.
Werner, L. A. (**1995**). ''Observer-based approaches to human infant psychoacoustics,'' in *Methods in Comparative Psychoacoustics*, edited by G. M. Klump, R. J. Dooling, R. R. Fay, and W. C. Stebbins (Birkhäuser, Boston), pp. 135–146.
Werner, L. A., and Bargones, J. Y. (**1991**). ''Sources of auditory masking in infants: Distraction effects,'' Percept. Psychophys. **50**, 405–412.
Werner, L. A., Folsom, R. C., and Mancl, L. R. (**1993**). ''The relationship between auditory brainstem response and behavioral thresholds in normal hearing infants and adults,'' Hearing Res. **68**, 131–141.
Werner, L. A., Folsom, R. C., and Mancl, L. R. (**1994**). ''The relationship between auditory brainstem response latency and behavioral thresholds in normal hearing infants and adults,'' Hearing Res. **77**, 88–98.
Werner, L. A., Kopyar, B. A., and Bargones, J. Y. (**1993**). ''Psychophysical measures of auditory sensitivity in human infants,'' Abstracts Midwinter Research Meeting ARO **16**, 53.
Werner, L. A., and Marean, G. C. (**1991**). ''Methods for estimating infant thresholds,'' J. Acoust. Soc. Am. **90**, 1867–1875.
Werner, L. A., Marean, G. C., Halpin, C. F., Spetner, N. B., and Gillenwater, J. M. (**1992**). ''Infant auditory temporal acuity: Gap detection,'' Child Dev. **63**, 260–272.
Wightman, F., and Allen, P. (**1992**). ''Individual differences in auditory capability among preschool children,'' in *Developmental Psychoacoustics*, edited by L. A. Werner and E. W. Rubel (American Psychological Association, Washington, DC), pp. 113–133.
Wightman, F., Allen, P., Dolan, T., Kistler, D., and Jamieson, D. (**1989**). ''Temporal resolution in children,'' Child Dev. **60**, 611–624.

Psychometric functions for discrimination of two-component complex tones in listeners with normal hearing and listeners with hearing loss

Kathryn Hoberg Arehart and Peninah Fine Rosengard
University of Colorado at Boulder, Speech, Language and Hearing Sciences, Campus Box 409, Boulder, Colorado 80309

(Received 7 August 1997; revised 2 April 1998; accepted 17 December 1998)

This study compared the ability of 5 listeners with normal hearing and 12 listeners with moderate to moderately severe sensorineural hearing loss to discriminate complementary two-component complex tones (TCCTs). The TCCTs consist of two pure tone components (f_1 and f_2) which differ in frequency by Δf (Hz) and in level by ΔL (dB). In one of the complementary tones, the level of the component f_1 is greater than the level of component f_2 by the increment ΔL; in the other tone, the level of component f_2 exceeds that of component f_1 by ΔL. Five stimulus conditions were included in this study: $f_c=1000$ Hz, $\Delta L=3$ dB; $f_c=1000$ Hz, $\Delta L=1$ dB; $f_c=2000$ Hz, $\Delta L=3$ dB; $f_c=2000$ Hz, $\Delta L=1$ dB; and $f_c=4000$ Hz, $\Delta L=3$ dB. In listeners with normal hearing, discrimination of complementary TCCTs (with a fixed ΔL and a variable Δf) is described by an inverted U-shaped psychometric function in which discrimination improves as Δf increases, is (nearly) perfect for a range of Δf's, and then decreases again as Δf increases. In contrast, group psychometric functions for listeners with hearing loss are shifted to the right such that above chance performance occurs at larger values of Δf than in listeners with normal hearing. Group psychometric functions for listeners with hearing loss do not show a decrease in performance at the largest values of Δf included in this study. Decreased TCCT discrimination is evident when listeners with hearing loss are compared to listeners with normal hearing at both equal SPLs and at equal sensation levels. In both groups of listeners, TCCT discrimination is significantly worse at high center frequencies. Results from normal-hearing listeners are generally consistent with a temporal model of TCCT discrimination. Listeners with hearing loss may have deficits in using phase locking in the TCCT discrimination task and so may rely more on place cues in TCCT discrimination. © *1999 Acoustical Society of America.* [S0001-4966(99)00604-9]

PACS numbers: 43.66.Fe, 43.66.Hg, 43.66.Mk [JWH]

INTRODUCTION

Listeners with hearing loss show deficits in frequency discrimination tasks using pulsed pure tones (e.g., Turner and Nelson, 1982; Freyman and Nelson, 1991; Moore and Peters, 1992; Simon and Yund, 1993), frequency-modulated signals (e.g., Zurek and Formby, 1981; Grant, 1987), and steady-state complex tones (e.g., Moore and Peters, 1992; Arehart, 1994). These deficits have been attributed to problems both in place coding and in temporal coding (for a review, see Moore, 1995). The purpose of this paper is to add to our understanding of frequency discrimination in listeners with hearing loss by measuring their ability to discriminate complementary pairs of two-component complex tones (TCCTs).

TCCTs have a time-invariant long-term spectrum, which can be described in terms of two pure tone components (f_1 and f_2) which differ in frequency by Δf (Hz) and in level by ΔL (dB). Figure 1 shows a pair of complementary TCCT tones. In one of the complementary tones, the level of the component f_1 is greater than the level of component f_2 by the increment ΔL (Fig. 1, bottom panel), whereas in the other complementary tone, the level of component f_2 exceeds that of component f_1 by ΔL (Fig. 1, middle panel).

In listeners with normal hearing, discrimination of complementary TCCT pairs (with a fixed ΔL and a variable Δf) is described by an inverted U-shaped psychometric function in which discrimination ability improves as Δf increases, is (nearly) perfect for a range of Δf's, and then decreases again as Δf increases (Feth, 1974; Feth and O'Malley, 1977). TCCT discrimination could be based on either a place or on a temporal mechanism.

The discrimination of the TCCT pairs shown by the rising portions of the psychometric function has been explained by models using the weighted average instantaneous frequency of TCCTs. The envelope weighted average instantaneous frequency (EWAIF) and intensity weighted average instantaneous frequency (IWAIF) models (Feth, 1974; Anantharaman *et al.*, 1993) predict that when the components f_1 and f_2 are of equal level (i.e., $\Delta L=0$ dB), the pitch of a TCCT (labeled P_c in Fig. 1, top panel) corresponds to the center frequency (f_c). When ΔL is not 0 dB, the pitch of the signal shifts towards that of the more intense component. That is, when f_2 is greater in level than f_1, then the TCCT will have a pitch corresponding to a frequency above f_c (labeled P_2 in Fig. 1, mid panel); conversely, when f_1 is greater in level than f_2, the TCCT will have a pitch corresponding to a frequency below f_c (labeled P_1 in Fig. 1, bottom panel). According to the EWAIF/IWAIF models, a listener will be able to discriminate complementary TCCTs if the predicted

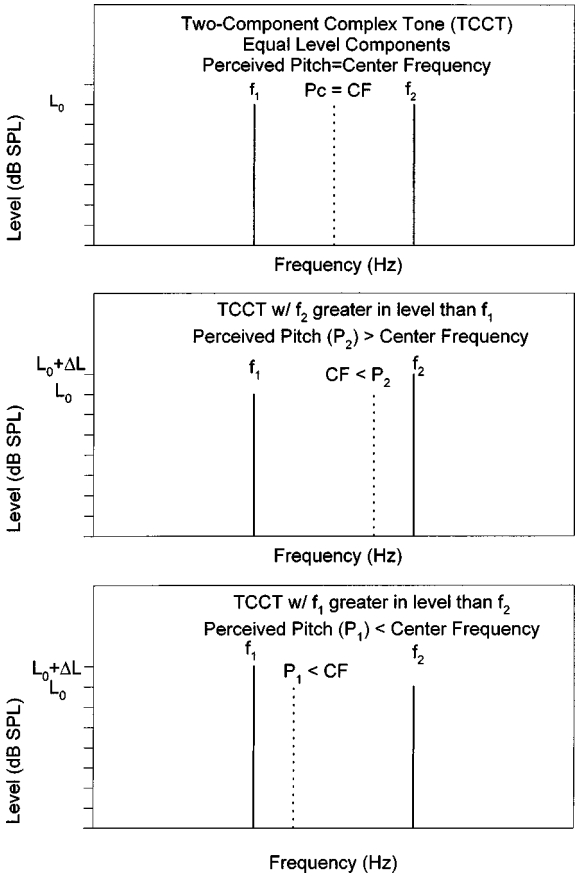


FIG. 1. Long-term amplitude spectrum of a two-component complex tone (TCCT) consists of two pure tone components (f_1 and f_2) which differ in frequency by Δf (Hz) and in level by ΔL (dB). Top panel shows components of f_1 and f_2 with equal level. Middle panel shows f_2 greater in level than f_1. Bottom panel shows f_1 greater in level than f_2.

pitch difference (P_2-P_1) is equal to or exceeds the listener's pure tone frequency difference limen. Feth *et al.* (1982) used a forced choice discrimination task to measure the ability of normal-hearing listeners to discriminate pairs of complementary TCCTs. Listeners were also asked to match a complex tone containing two equal-amplitude components to each member of a TCCT pair. The difference in the perceived pitches of complementary TCCTs increased as their discriminability increased.

Feth and O'Malley (1977) have suggested that the point on the psychometric function at which performance worsens at large Δf's is related to auditory spectral resolution. When the frequency difference between the two tones exceeds the critical band, the listener presumably hears two resolved tones rather than the unitary pitch evoked by unresolved components in a TCCT signal (e.g., Plomp, 1964; Plomp and Mimpen, 1968). The task then becomes more difficult for the listener because the frequencies of the resolved tones are the same on both intervals of a trial.

No published data are currently available regarding TCCT discrimination in listeners with hearing loss. Because sensorineural hearing loss is often accompanied by spectro-temporal processing deficits (for a review, see Moore, 1995), the shape of psychometric functions for TCCT discrimination may differ in listeners with hearing loss. Deficits in either place and/or temporal mechanisms might result in abnormally large difference limens for TCCT discrimination. Because sensorineural hearing loss is often characterized by abnormally wide auditory filters, the unitary pitch of TCCT tones might not break down at the same point that it does in listeners with normal hearing. The present study had two primary goals: (1) to compare the shape of psychometric functions for TCCT discrimination in listeners with normal hearing and listeners with hearing loss at center frequencies of 1000, 2000, and 4000 Hz; and (2) to evaluate the extent to which TCCT discrimination can be explained by temporal and/or place mechanisms.

TABLE I. Age and audiometric thresholds (in dB HL) of listeners with hearing loss.

		Frequency (Hz)					
Listener	Age	250	500	1000	2000	3000[a]	4000
I1	64	20	20	30	35[b]	40	35
I2	67	10	15	25	35[b]		45[b]
I3	69	45	50	60[b]	55		50[b]
I4	36	35	50	50[b]	60[b]		60[b]
I5	38	25	35	50[b]	55		60
I6	62	65	65	50[b]	55	60	70
I7	65	15	15	15	55[b]	50	55[b]
I8	54	35	20	15	55[b]	65	65[b]
I9	65	25	50	65[b]	55		65
I10	70	10	0	0	30	50	60[b]
I11	62	10	10	20	45[b]		60
I12	67	25	25	30	55[b]	65	70[b]

[a]Thresholds were not available for all listeners at 3000 Hz.
[b]Indicates that listener was tested at this frequency.

I. METHOD

A. Listeners

Participants in this study included 5 listeners with normal hearing (age range 22–32 years) and 12 listeners with moderate to moderately severe sensorineural hearing loss of presumably cochlear origin (age range 36–70 years). Normal hearing is defined here as thresholds of 15 dB HL (ANSI, 1989) or better at octave frequencies from 250–8000 Hz, inclusive. A moderate to moderately severe hearing loss refers to thresholds of at least 35 dB HL but not exceeding 70 dB HL, for at least one of the following frequencies: 1000, 2000, or 4000 Hz. All listeners were tested monaurally and individually in a double-walled sound-treated booth. Because of time availability and/or configuration of hearing loss, not all listeners were tested at all three center frequencies (1000, 2000, and 4000 Hz). Daily test sessions were usually 1 and limited to 2 h. Listeners were paid $6/h for their participation.

During their initial visit, listeners underwent an audiometric evaluation. All of the listeners with hearing loss demonstrated test results that were consistent with cochlear impairment: normal tympanometry; absence of excessive acoustic reflex decay; and absence of air–bone gap exceeding 10 dB at any frequency. Table I provides a summary of the audiometric thresholds of the listeners with hearing loss and indicates at which frequencies TCCT discrimination was measured in individual listeners.

B. Stimuli

Individual components of the two-component complex tones were digitally generated with two 16-bit digital-to-analog (D-A) converters at a sampling rate of 25 000 Hz. The tones were characterized by a 500-ms duration, 0 degree starting phase, and 20-ms rise–fall times shaped by a raised cosine function. The delivery of the stimuli was controlled by a microprocessor which interfaced with signal conditioning equipment and a subject response box. The stimuli were played out in the following way: the two components of a TCCT signal were each routed separately through low-pass filters (cutoff frequency=10 000 Hz) and programmable attenuators, then were combined in a mixer, and finally delivered to the listener's ear through either an Etymotic ER2 insert earphone or a TDH-49-P headphone. All listeners were tested with ER2 earphones except for listener I8 (f_c =2 kHz, ΔL=3 dB; f_c=2 kHz, ΔL=1 dB; f_c=4 kHz, ΔL=3 dB) and for listener I4 (f_c=4 kHz, ΔL=3 dB).

Five stimulus conditions were included in this study: f_c=1000 Hz, ΔL=3 dB; f_c=1000 Hz, ΔL=1 dB; f_c =2000 Hz, ΔL=3 dB; f_c=2000 Hz, ΔL=1 dB; and f_c =4000 Hz, ΔL=3 dB. For each of these stimulus conditions, complementary pairs of TCCTs were synthesized for Δf=2.5, 5, 10, 20, 50, 100, 200, 283, 500, 707, and 1000 Hz. Level differences (ΔL=1 dB; ΔL=3 dB) between the stimulus components were determined at the output of the earphones, using a Zwislocki coupler for the insert earphones and a 6-cc coupler for the TDH-49 earphones. The overall level of the TCCTs for listeners with hearing loss ranged from 75–90 dB SPL, with the requirement that the signals be at least 20 dB sensation level (SL) for the listeners. To facilitate comparisons with listeners with hearing loss at both equal sound pressure level and at equal sensation level, listeners with normal hearing were tested at 75 dB SPL for all five stimulus conditions and at 25 dB SL for f_c=1000 Hz, ΔL=1 dB and f_c=2000 Hz, ΔL=3 dB. Five listeners with normal hearing (N1, N2, N3, N4, and N5) were tested at all three center frequencies. Five listeners with hearing loss (I3, I4, I5, I6, and I9) were tested at 1000 Hz, seven listeners with hearing loss were tested at 2000 Hz (I1, I2, I4, I7, I8, I11, and I12) and seven listeners with hearing loss were tested at 4000 Hz (I2, I3, I4, I7, I8, I10, and I12).

C. Procedure

Discrimination of complementary TCCTs was measured in a two-alternative forced-choice procedure with feedback at f_c=1000, 2000, and 4000 Hz. On each trial, listeners heard both members of a TCCT pair, presented in random order. Listeners were asked to pick the interval containing the signal with the higher pitch. Frequency separation, level difference, and center frequency were fixed for blocks of 50 trials. In each block, a listener's performance was measured by the percentage of trials in which the TCCT with the higher-level high frequency component (Fig. 1, middle panel) was identified as being the higher-pitched signal. In each test session, a single psychometric function based on nine blocks (Δf=10 Hz through 1000 Hz) of 50 trials was measured (i.e., f_c and ΔL were fixed across blocks; each value of Δf was tested once, in random order). Listeners whose performance at Δf=10 Hz exceeded d'=1 were also tested at Δf=2.5 Hz and Δf=5 Hz. At least six blocks (300 trials) were obtained for each listener in each $f_c/\Delta L/\Delta f$ condition. The first 100 trials were considered practice. The data reported here for each listener in each $f_c/\Delta L/\Delta f$ condition are based on the remaining 200 trials.

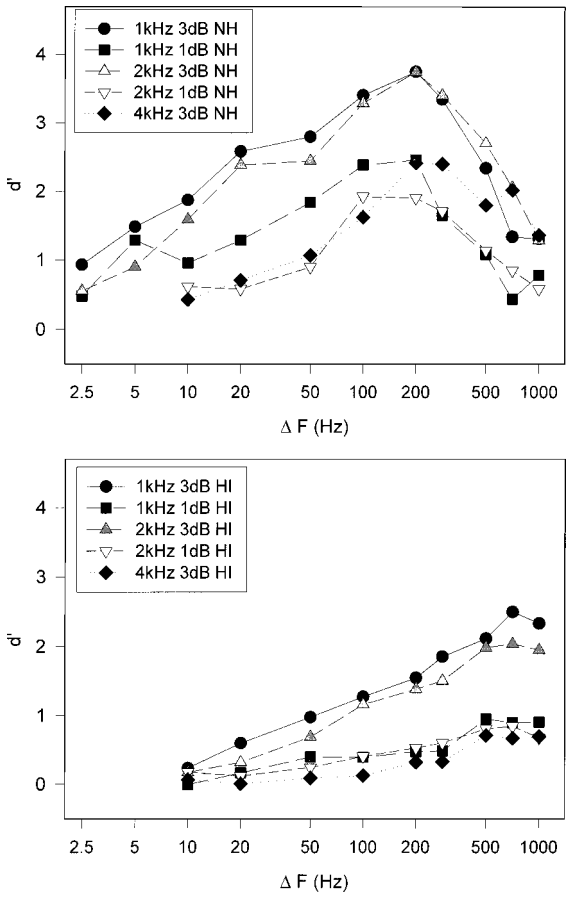


FIG. 2. Performance on the TCCT discrimination task is plotted as a function of log Δf (Hz). Discrimination of TCCTs is expressed in terms of d'. Following Macmillan and Creelman (1991),[1] perfect discrimination scores were converted to 99.75%, which corresponds to a d'=4, shown parenthetically on the right vertical axis. Psychometric functions are shown for listeners with normal hearing (top panel) and for listeners with hearing loss (bottom panel) at f_c=1000, 2000, and 4000 Hz. Psychometric functions are based on mean performance of five listeners with normal hearing and five to seven listeners with hearing loss.

II. RESULTS AND DISCUSSION

A. Group psychometric functions

Percent correct values for each listener in each $f_c/\Delta L/\Delta f$ condition were converted to d' values (Hacker and Ratcliff, 1979).[1] Discrimination performance (expressed in terms of d') was then plotted as a function of log Δf (Figs. 2–5). Figure 2 shows discrimination averaged across listeners with normal hearing (top panel) and averaged across listeners with hearing loss (bottom panel) for the five $f_c/\Delta L$ conditions. Functions for listeners with normal hearing were obtained at 75 dB SPL. Functions for listeners with hearing loss were obtained at levels ranging from 75–90 dB SPL.

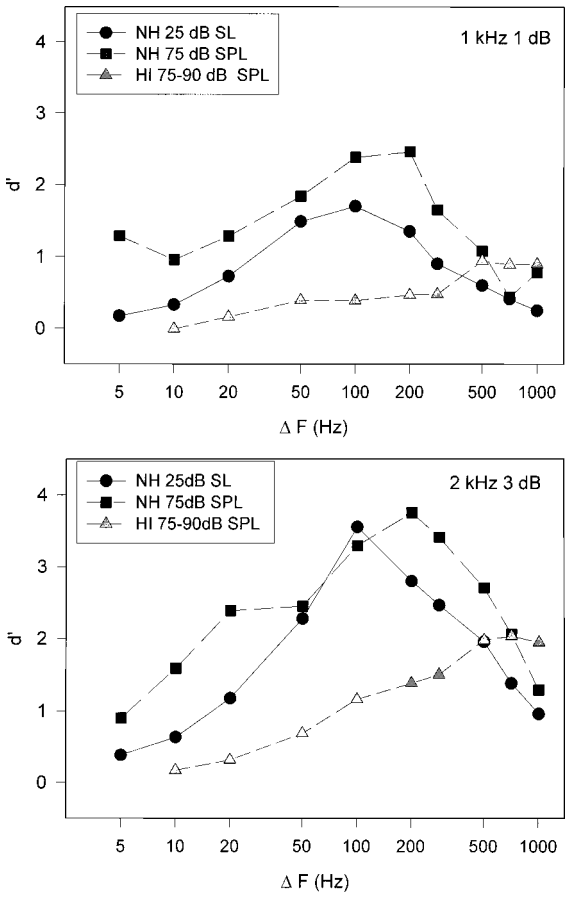


FIG. 3. Plotted as in Fig. 2, psychometric functions for TCCT discrimination based on average performance by listeners with normal hearing and listeners with hearing loss are shown for two conditions in which listeners with normal hearing were tested at 75 and 25 dB SL ($f_c = 1000$ Hz, $\Delta L = 1$ dB, top panel; $f_c = 2000$ Hz, $\Delta L = 3$ dB, bottom panel). Psychometric functions for listeners with hearing loss are also shown.

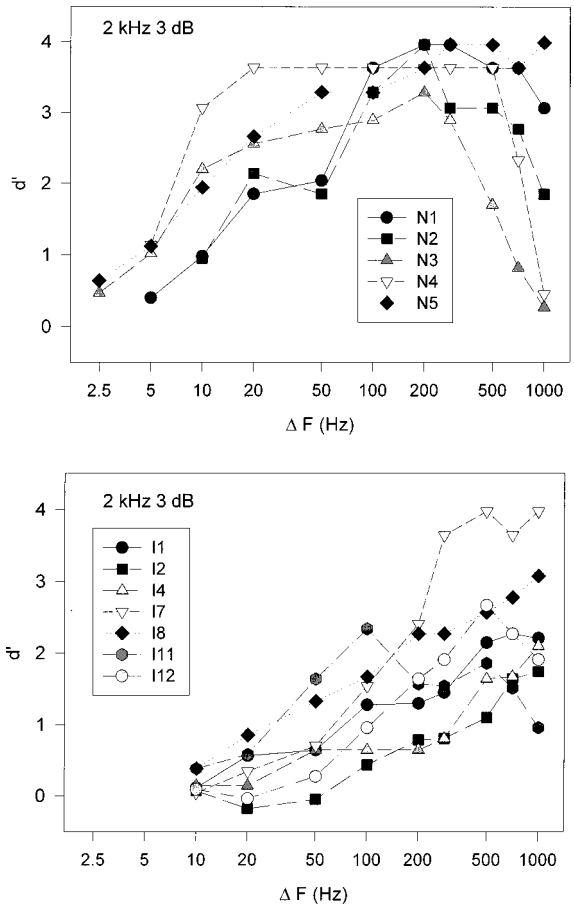


FIG. 4. Plotted as in Fig. 2, psychometric functions for TCCT discrimination are shown for individual listeners with normal hearing (top panel) and with hearing loss (bottom panel) in the $f_c = 2000$ Hz, $\Delta L = 3$ dB condition.

Consistent with reports in the literature (Feth, 1974; Feth and O'Malley, 1977), average TCCT discrimination for the group of listeners with normal hearing is described by inverted U-shaped psychometric functions in which discrimination improves as Δf increases, is (nearly) perfect for a range of Δf's, and then decreases again as Δf increases. In contrast, group psychometric functions for listeners with hearing loss are shifted to the right such that above-chance performance occurs at larger values of Δf than in listeners with normal hearing. Furthermore, group psychometric functions for listeners with hearing loss do not show a decrease in performance at the largest values of Δf included in this study.

These differences in psychometric function shape between listeners with normal hearing and listeners with hearing loss are evident in both equal SPL and equal SL comparisons. Figure 3 shows mean performance as a function of log Δf for the group of listeners with normal hearing (at both 75 dB and 25 dB SL) and for the group of listeners with hearing loss (at 75–90 dB SPL) for $f_c = 1000$ Hz, $\Delta L = 1$ dB (top panel) and $f_c = 2000$ Hz, $\Delta L = 3$ dB (bottom panel). While performance of normal-hearing listeners is worse in the 25 dB SL condition, both the 75 dB SPL and 25 dB SL functions from listeners with normal hearing show the characteristic inverted U-shaped function.

B. Individual psychometric functions

Some intersubject variability was evident in both listener groups. For example, Figs. 4 and 5 show psychometric functions for individual listeners with normal hearing and with hearing loss tested at $f_c = 2000$ Hz, $\Delta f = 3$ dB and $f_c = 1000$ Hz, $\Delta L = 1$ dB. Psychometric functions for individual listeners with hearing loss vary in terms of how far they are shifted to the right. Variability is evident among normal-hearing listeners in both the rising and falling portions of the psychometric functions. For example, some listeners maintain excellent discrimination at large Δf's while others show decreasing ability to discriminate TCCT tones for values of Δf greater than 200 Hz. These individual differences are addressed in more detail below.

C. Rising part of psychometric functions

1. Difference limens for TCCT discrimination

The difference limens (DLs) for TCCT discrimination were defined as the Δf corresponding to $d' = 1$ (76%) on the rising part of the psychometric function. Using this definition, comparisons are readily made between conditions and groups as well as between predictions made by place and

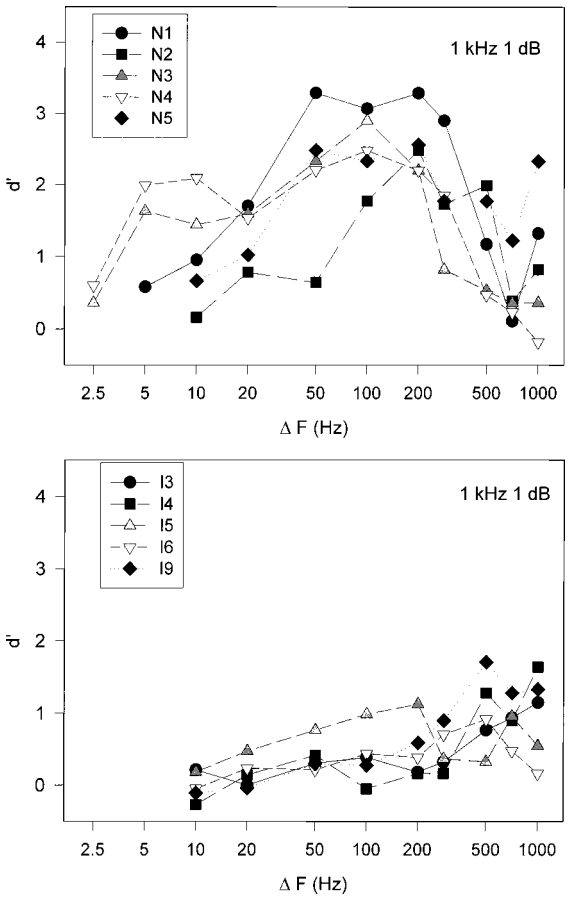


FIG. 5. Plotted as in Fig. 2, psychometric functions for TCCT discrimination are shown for individual listeners with normal hearing (top panel) and with hearing loss (bottom panel) in the $f_c = 1000$ Hz, $\Delta L = 1$ dB condition.

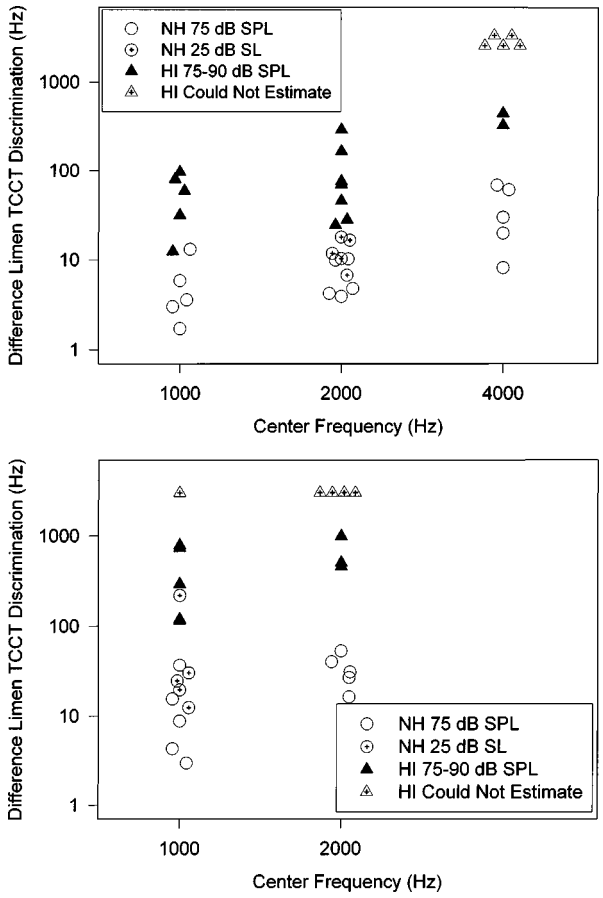


FIG. 6. Difference limens (DLs) for TCCT discrimination in normal-hearing (NH) and hearing-impaired (HI) listeners are shown as a function of center frequency for $\Delta L = 3$ dB (top panel) and $\Delta L = 1$ dB (bottom panel). The DLs for several listeners could not be estimated due to discrimination which never reached $d' = 1$.

temporal models. For each listener in each condition, discrimination (expressed in terms of d') was plotted as a function of $\log \Delta f$. Each d'-$\log \Delta f$ function was visually inspected to determine the points belonging to the rising portion of the function. Generally, the points belonging to the rising portion of psychometric functions included values of d' ranging from 0.18 (55%) to 2.33 (95%). The rising portions of the d'-$\log \Delta f$ functions were then fit with least squares linear regression lines. The mean coefficient of determination (R^2) for these linear regression fits for individual listeners was 0.89. Finally, the Δf corresponding to $d' = 1$ (76%) was determined by the equation for the linear regression line.

The DLs for TCCT discrimination for individual listeners are shown as a function of center frequency for $\Delta L = 3$ dB (Fig. 6, top panel) and $\Delta L = 1$ dB (Fig. 6, bottom panel). The DLs for TCCT discrimination could not be estimated for several listeners with hearing loss whose performance did not reach $d' = 1$. This includes one listener at $f_c = 1000$ Hz, $\Delta L = 1$ dB, four listeners at $f_c = 2000$ Hz, $\Delta L = 1$ dB, and five listeners at $f_c = 4000$ Hz, $\Delta L = 3$ dB.

Nonparametric analysis (the Mann Whitney test for two independent samples) was used to assess whether statistically significant differences existed between the DLs for TCCT discrimination for listeners with normal hearing and listeners with hearing loss in each condition. Between group comparisons made at high overall levels (Fig. 2) show significant differences between the two groups for the DLs at $f_c = 1000$ Hz, $\Delta L = 3$ dB ($p = 0.016$); at $f_c = 1000$ Hz, $\Delta L = 1$ dB ($p = 0.016$); at $f_c = 2000$ Hz, $\Delta L = 3$ dB ($p = 0.003$); and at $f_c = 2000$ Hz, $\Delta L = 1$ dB ($p = 0.036$). No significant differences existed between the two groups for the DLs for TCCT discrimination at $f_c = 4000$ Hz, $\Delta L = 3$ dB ($p = 0.095$). However, the statistical analysis at $f_c = 4000$ Hz, $\Delta L = 3$ dB is limited by the fact that only two listeners with hearing loss had DLs for TCCT discrimination that could be estimated in this condition.

Significant differences also existed between the normal-hearing and hearing-loss groups in the two conditions in which equal sensation level comparisons were made (Fig. 3): $f_c = 1000$ Hz, $\Delta L = 1$ dB ($p = 0.032$) and $f_c = 2000$ Hz, $\Delta L = 3$ dB ($p = 0.003$).

The Wilcoxon signed ranks test for two related samples was used to assess whether statistically significant differences existed across frequency. While there were several listeners with hearing loss tested at more than one frequency, there was an insufficient number of listeners with hearing loss who had DLs that could be estimated at more than one frequency. Therefore, this analysis included data only from listeners with normal hearing. There were no significant differences between $f_c = 1000$ Hz, $\Delta L = 3$ dB and $f_c = 2000$ Hz, $\Delta L = 3$ dB ($p = 0.138$). There were significant

differences between $f_c = 1000$ Hz, $\Delta L = 1$ dB and $f_c = 2000$ Hz, $\Delta L = 1$ dB ($p = 0.043$); between $f_c = 1000$ Hz, $\Delta L = 3$ dB and $f_c = 4000$ Hz, $\Delta L = 3$ dB ($p = 0.043$); and between $f_c = 2000$ Hz, $\Delta L = 3$ dB and $f_c = 4000$ Hz, $\Delta L = 3$ dB ($p = 0.043$). While not amenable to statistical tests, the trend for performance to worsen at high center frequencies was also evident in listeners with hearing loss: Average DLs for TCCT discrimination determined from the group psychometric functions for listeners with hearing loss were 51 Hz for $f_c = 1000$ Hz, $\Delta L = 3$ dB and 78 Hz for $f_c = 2000$ Hz, $\Delta L = 3$ dB. The DLs for TCCT discrimination were 484 and 331 Hz for the two listeners with estimable DLs at $f_c = 4000$ Hz, $\Delta L = 3$ dB. (Five of the seven listeners with hearing loss tested at f_c had performance which did not reach $d' = 1$ at 4000 Hz.)

The Wilcoxon signed ranks test was also used to assess whether statistically significant differences existed between $\Delta L = 3$ dB and $\Delta L = 1$ dB. Only subjects who had estimable difference limens at $f_c = 1000$ Hz, $\Delta L = 3$ dB and $f_c = 1000$ Hz, $\Delta L = 1$ dB or $f_c = 2000$ Hz, $\Delta L = 3$ dB and $f_c = 2000$ Hz, $\Delta L = 1$ dB were included. That is, all five listeners with normal hearing tested at 1000 Hz and 2000 Hz, four listeners with hearing loss tested at 1000 and three listeners with hearing loss tested at 2000 Hz were included. Performance was significantly better for $\Delta L = 3$ dB than for $\Delta L = 1$ dB at both 1000 Hz ($p = 0.008$) as well as at 2000 Hz ($p = 0.012$).

Finally, Wilcoxon signed ranks test for paired samples was also used to assess whether statistically significant differences existed between the equal-SL and equal-SPL conditions in the normal-hearing listeners. Results showed that there were statistically significant differences both at $f_c = 1$ kHz, $\Delta L = 1$ dB ($p = 0.043$) and at $f_c = 2$ kHz, $\Delta L = 3$ dB ($p = 0.043$).

In summary, difference limens for TCCT discrimination were significantly larger for listeners with hearing loss than for listeners with normal hearing. This decreased discrimination ability is evident when listeners with hearing loss are compared to listeners with normal hearing both at equal SPLs and at equal sensation levels. Furthermore, in both listener groups, TCCT discrimination is significantly worse at 4000 Hz than at lower center frequencies and is worse for $\Delta L = 1$ dB than for $\Delta L = 3$ dB.

Discrimination of TCCTs could be based on either a place mechanism or a temporal mechanism. The following two sections consider the extent to which discrimination in the two groups of listeners can be explained by an excitation pattern model and/or by a model based on the temporal encoding of weighted average instantaneous frequencies.

2. Evaluation of an excitation pattern model

If TCCT discrimination is based on differences in the excitation patterns of complementary TCCTs, then listeners with hearing loss may have deficits in TCCT discrimination due to decreased frequency selectivity associated with cochlear hearing loss (e.g., Tyler *et al.*, 1982; Carney and Nelson, 1983). Current evidence supports the idea that deficits in frequency selectivity contribute to abnormally large frequency modulation difference limens (Moore, 1995).

To examine the possible role of place information in TCCT discrimination, we considered two measures of the differences in excitation level between the two excitation patterns of members of a TCCT pair. The first measure, the summed excitation pattern difference (ΔL_{sum}), assumes that listeners combine information across different regions of the excitation pattern in their discrimination of TCCT pairs (e.g., Florentine and Buus, 1981; Moore and Sek, 1994). The second measure, the maximum excitation pattern difference (ΔL_{max}), is based on the idea that listeners base their discriminations on the one location on the excitation pattern that shows the greatest change in level.

Excitation patterns for TCCT stimuli were calculated for both normal and impaired auditory systems using the model of Moore and Glasberg (1997). This model is similar to previously published models in that the excitation pattern for a sound is determined by the output of auditory filters in response to that sound (Glasberg and Moore, 1990). Outer hair cell damage is modeled as an increase in the width of the auditory filters and a concomitant broadening of the excitation pattern. Details regarding the calculations in the model are described in the Appendix.

In implementing the Moore and Glasberg model, the following assumptions were made. First, the normal auditory system was assumed to have 0 dB HL thresholds at all frequencies and the level of the lower intensity pure tone component in each TCCT signal was 75 dB SPL. The parameters used for the impaired auditory system differed slightly for each center frequency. Specifically, excitation patterns for the impaired auditory system at a particular center frequency were based on thresholds averaged across the subset of hearing-impaired listeners who were tested at that test frequency. The level of the lower intensity pure tone component in each TCCT signal was the average of the stimulus levels used for hearing-impaired listeners tested at each center frequency (81 dB SPL at $f_c = 1000$ Hz and 84 dB SPL at $f_c = 2000$ Hz). The proportion of hearing loss at each audiometric frequency that was attributable to outer hair cell loss was assumed to be 80%, which is the default value in the model. Finally, the excitation pattern analysis does not include data from hearing-impaired listeners at $f_c = 4000$ Hz since the group of listeners with hearing loss was essentially unable to discriminate TCCT signals at this frequency.

Figure 7 (top panel) shows excitation patterns for complementary TCCT pairs corresponding to the group difference limens for TCCT discrimination at $f_c = 2000$ Hz, $\Delta L = 3$ dB. The DL for normal-hearing listeners corresponded to 4.8 Hz; the DL for hearing-impaired listeners corresponded to 78.7 Hz. Excitation level is plotted on a scale in which frequency has been transformed to equivalent rectangular bandwidth (ERB) number.[2] The differences between these excitation patterns (Fig. 7, bottom panel) are very small in the normal auditory system but as large as 1.8 dB on the low-frequency side of the excitation patterns in the impaired auditory system.

Differences between excitation patterns for TCCT signals were quantified in the following way. First, excitation pattern differences were evaluated for all channels (n) showing suprathreshold levels. (The absolute number of suprath-

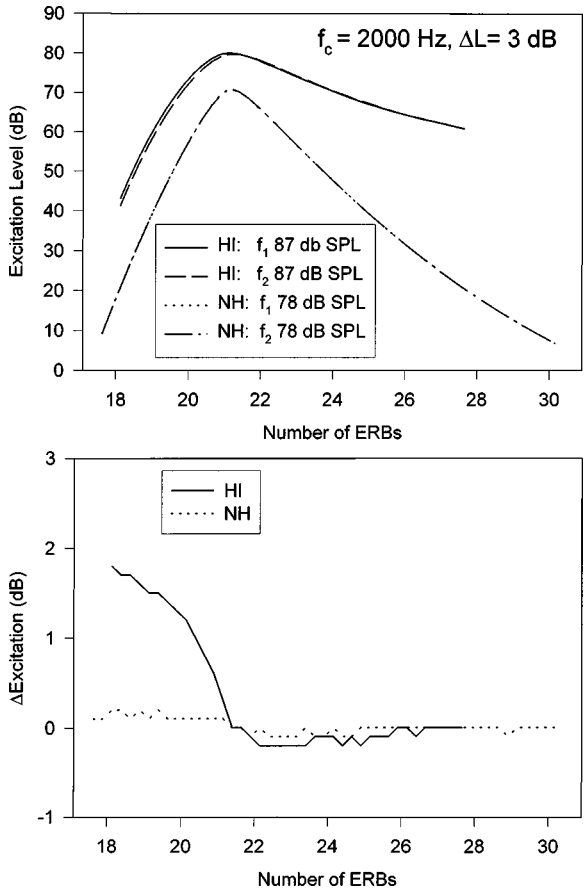


FIG. 7. Top panel shows excitation patterns for complementary TCCT pairs corresponding to the group difference limens for TCCT discrimination at $f_c = 2000$ Hz, $\Delta L = 3$ dB. Excitation level is plotted as a function of frequency, expressed in terms of the equivalent rectangular bandwidth scale (Glasberg and Moore, 1990).[2] Two excitation patterns are shown for a normal auditory system for a pair of TCCTs with a $\Delta f = 4.8$ Hz. In one TCCT, the level of f_1 is 78 dB SPL and of f_2 is 75 dB SPL; for the other TCCT, the level of f_1 is 75 dB SPL and for f_2 is 78 dB SPL. Excitation patterns for the impaired auditory system are for a pair of TCCTs with $\Delta f = 78.7$ Hz. In one TCCT, the level for f_1 is 87 dB SPL and for f_2 is 84 dB SPL; for the other TCCT, the level for f_1 is 84 dB SPL and for f_2 is 87 dB. The bottom panel shows the differences in excitation between the two excitations patterns for the impaired auditory system and the two excitation patterns for the normal auditory system.

reshold channels was very similar across all TCCT pairs.) Each channel was defined in terms of one ERB (Glasberg and Moore, 1990). Second, the summed difference was calculated as $\sqrt{(\Sigma \Delta L_i^2)}$, where ΔL_i (in dB) are the excitation level differences measured in a total of n 1-ERB intervals. Finally, $\Delta L_{\max}$ was defined as the greatest difference in level in a channel along the excitation patterns.

The overall differences and maximum differences in the active (suprathreshold) channels in excitation patterns increase as the discriminability of TCCT pairs increases. By way of example, Table II shows ΔL_{sum} and $\Delta L_{\max}$ for excitation patterns calculated for normal and impaired auditory systems at $f_c = 1000$ Hz. The trend for excitation pattern differences to increase as discriminability increases was confirmed by least square linear regression fits which showed strong predictive relationship between ΔL_{sum} and d' and $\Delta L_{\max}$ and d' at all three center frequencies in normal-hearing listeners (mean $R^2 = 0.96$) and in listeners with hearing loss at 1000 and 2000 Hz (mean $R^2 = 0.85$).

However, the excitation pattern differences for complementary TCCT pairs corresponding to group DLs for TCCT discrimination (listed in Table III) challenge the plausibility of the excitation pattern model being able to account for TCCT discrimination in normal-hearing listeners. In order for TCCT discrimination to be consistent with an excitation pattern model, values of ΔL_{sum} and $\Delta L_{\max}$ for TCCTs corresponding to the DL ($d' = 1$) should be similar across listener groups and across conditions within a group (Moore, 1998). Here ΔL_{sum} and $\Delta L_{\max}$ differ between the normal-hearing and hearing-impaired groups; ΔL_{sum} and $\Delta L_{\max}$ also differ across conditions within the normal-hearing group. Furthermore, if we assume, as Zwicker (1970) has, that a 1-dB difference in excitation level is required for a listener to detect frequency differences based on place information, then the values of $\Delta L_{\max}$ for the normal auditory system are much too small to account for TCCT discrimination. This conclusion is consistent with other studies. Versfeld and Houtsma (1995) reported that excitation pattern differences were too small to account for the results of their experiment in which they used an adaptive procedure to determine the level difference (ΔL) required to just discriminate TCCTs with a fixed Δf. Stover and Feth (1983) also reported that differences in the center of gravity of excitation patterns were too small to account for TCCT discrimination in normal-hearing listeners.

In contrast, TCCT discrimination in the hearing-impaired listeners is consistent with the excitation pattern model. As shown in Table III, the excitation pattern differences calculated for an impaired auditory system are relatively constant across conditions and exceed the criterion value of 1 dB. The predictions shown in Table III are based on group psychometric functions which represent data averaged across the hearing-impaired listeners. Similar predictions of the excitation pattern model were obtained for excitation patterns based on psychometric functions from individual hearing-impaired listeners.

3. Evaluation of the IWAIF model

Feth and colleagues (e.g., Burns and Feth, 1983; Stover and Feth, 1983) have suggested that TCCT discrimination in normal-hearing listeners is based on temporal encoding of the weighted average instantaneous frequency. The finding that TCCT discrimination is significantly worse at $f_c = 4000$ Hz is consistent with physiological data showing that phase locking deteriorates at higher frequencies and disappears above 5000 Hz (Rose *et al.*, 1967). Several investigators have suggested that reduced pure tone frequency discrimination in listeners with hearing loss is due to reduced ability to use frequency information conveyed by phase-locked responses (e.g., Wakefield and Nelson, 1985). Similarly, deficits in neural synchrony might reduce the precision with which listeners with hearing loss could encode the average instantaneous frequencies of TCCT signals. However, physiological data in support of the idea that cochlear hearing loss results in deficits in neural phase-locking is limited (Harrison and Evans, 1979; Woolf *et al.*, 1981).

TABLE II. Differences predicted by the IWAIF model (ΔIWAIF), summed excitation pattern difference (ΔL_{sum}), and maximum excitation pattern difference (ΔL_{max}) are listed for complementary TCCT pairs shown on the rising part of the group psychometric functions at $f_c = 1000$ for normal-hearing (NH) and hearing-impaired (HI) listeners. The mean percent correct scores obtained in the 2AFC task and associated d' values are also shown.

Condition	Group	ΔF (Hz)	% Correct	d'	ΔIWAIF (Hz)	ΔL_{max} (dB)	ΔL_{sum} (dB)
$f_c = 1$ kHz, $\Delta L = 3$ dB	NH	2.5	75	0.93	0.8	0.10	0.21
	NH	5	85	1.48	1.6	0.30	0.60
	NH	10	91	1.87	3.3	0.70	1.23
	NH	20	97	2.58	6.7	1.20	2.27
	NH	50	98	2.80	16.7	2.30	4.91
	NH	100	99	3.41	33.4	2.90	7.92
$f_c = 1$ kHz, $\Delta L = 1$ dB	NH	10	75	0.95	1.2	0.20	0.35
	NH	20	82	1.28	2.3	0.40	0.68
	NH	50	90	1.84	5.7	0.80	1.72
	NH	100	95	2.38	11.5	1.00	2.69
$f_c = 1$ kHz, $\Delta L = 3$ dB	HI	10	57	0.23	3.3	0.40	0.54
	HI	20	66	0.59	6.7	0.70	1.07
	HI	50	75	0.97	16.6	1.60	2.55
	HI	100	82	1.27	33.2	2.60	4.50
	HI	200	86	1.54	66.5	3.00	6.41
	HI	283	90	1.85	94.0	3.00	8.06
	HI	500	93	2.11	166.1	3.10	9.96
$f_c = 1$ kHz, $\Delta L = 1$ dB	HI	20	55	0.16	2.3	0.30	0.36
	HI	50	61	0.39	5.7	0.60	0.84
	HI	100	61	0.39	11.5	0.90	1.47
	HI	200	63	0.47	22.9	1.00	2.19
	HI	283	63	0.48	32.4	1.00	2.64
	HI	500	75	0.94	57.3	1.10	3.25

We calculated the IWAIF for each TCCT signal (Feth, 1974; Anantharaman *et al.*, 1993), in order to determine the relationship between the weighted average instantaneous frequency of TCCT signals and the results obtained in our two groups of listeners. Because the predictions of the EWAIF and IWAIF models are very similar (Anantharaman *et al.*, 1993), the predictions from the IWAIF model are reported here due to the straightforward nature of the IWAIF calculations. The IWAIF model for the TCCTs used in this study can be described by the equation

$$\text{IWAIF} = (A_1 f_1 + A_2 f_2)/(A_1 + A_2)^2, \quad (1)$$

where A_1 and f_1 are, respectively, the amplitude and frequency of the lower frequency component, and A_2 and f_2 are, respectively, the amplitude and frequency of the higher frequency component.

We then determined the difference in the IWAIF values (ΔIWAIF) for each member of a TCCT pair (similar to $P_2 - P_1$ in Fig. 1). In Table II, ΔIWAIF values are shown together with average d' and percent correct values for complementary TCCTs shown on the rising part of group psychometric functions at $f_c = 1000$ Hz. Least squares linear regression fits showed strong relationship between d' and ΔIWAIF at all three center frequencies in listeners with normal hearing (mean $R^2 = 0.96$) and for $f_c = 1000$ Hz and $f_c = 2000$ Hz in listeners with hearing loss ($R^2 = 0.88$). (Data from listeners with hearing loss were not considered at $f_c = 4000$ Hz since the group of listeners with hearing loss was essentially unable to discriminate TCCT signals at this frequency.)

According to the EWAIF/IWAIF models, a listener will be able to discriminate complementary TCCTs if the predicted pitch difference ($P_2 - P_1$ in Fig. 1) is equal to or exceeds the listener's pure tone frequency difference limen. Table IV shows predicted IWAIF pitch differences corresponding to the DL for TCCT discrimination (as determined by a least squares linear regression line fit to the group psychometric functions and corresponding to $d' = 1$). Also shown are pure tone frequency difference limens (fdl's) reported by Wier *et al.* (1977), which correspond to 71% correct in a 2AFC task ($d' = 0.78$). While not directly compa-

TABLE III. Difference limens for TCCT discrimination (DL) and the maximum (ΔL_{max}) and summed (ΔL_{sum}) excitation pattern differences are shown for $d' = 1$, as estimated from the group psychometric functions for normal-hearing (NH) and hearing-impaired listeners (HI).

Condition	Group	DL (Hz)	ΔL_{max} (dB)	ΔL_{sum} (dB)
$f_c = 1$ kHz, $\Delta L = 3$ dB	NH	2.5	0.10	0.21
$f_c = 1$ kHz, $\Delta L = 1$ dB	NH	6.6	0.12	0.23
$f_c = 2$ kHz, $\Delta L = 3$ dB	NH	4.8	0.18	0.28
$f_c = 2$ kHz, $\Delta L = 1$ dB	NH	31.4	0.35	0.64
$f_c = 4$ kHz, $\Delta L = 3$ dB	NH	26.8	0.47	0.88
$f_c = 1$ kHz, $\Delta L = 1$ dB, 25 dB SL	NH	25.7	0.40	0.69
$f_c = 2$ kHz, $\Delta L = 3$ dB, 25 dB SL	NH	10.1	0.30	0.41
$f_c = 1$ kHz, $\Delta L = 3$ dB	HI	50.8	1.60	2.56
$f_c = 1$ kHz, $\Delta L = 1$ dB	HI	1964.0	1.30	4.50
$f_c = 2$ kHz, $\Delta L = 3$ dB	HI	78.7	1.80	2.41
$f_c = 2$ kHz, $\Delta L = 1$ dB	HI	2412.0	1.30	4.53

TABLE IV. Predicted IWAIF pitch differences (ΔIWAIF, in Hz) corresponding to the DLs for TCCT discrimination ($d'=1$), as estimated from the group psychometric functions for normal-hearing (NH) and hearing-impaired listeners (HI). ΔIWAIF values for the 25-dB sensation level condition tested in normal-hearing listeners are listed in parenthesis. Also shown are pure tone frequency differences limens reported by Wier, Jesteadt, and Green (1977) (WJG) for normal-hearing listeners ($d'=0.78$) for sensation levels of 20 and 80 dB.

	ΔIWAIF (Hz)					
Center frequency	NH $\Delta L=3$ dB	NH $\Delta L=1$ dB	HI $\Delta L=3$ dB	HI $\Delta L=1$ dB	WJG 80 dB SL	WJG 20 dB SL
1000	0.8	0.8 (3.0)	16.6	225.0	1.3	2.2
2000	1.6 (3.3)	3.6	26.2	276.5	2.3	5.8
4000	8.9	Not tested	Could not estimate	Not tested	11.4	21.8

rable due to differences in d', the ΔIWAIF values reported in Table IV are generally similar to the pure tone fdl's reported by Wier *et al.* (1977). Pure tone fdl's and ΔIWAIF values both decrease as sensation level increases and as center frequency decreases.

For a given center frequency, the ΔIWAIF values corresponding to the DL for TCCT discrimination should be constant for a given group of listeners. As shown in Table IV, ΔIWAIF values are the same (0.8 Hz) for normal-hearing listeners at 1000 Hz at 1 and 3 dB, but are not similar at 2000 Hz (values of 1.6 and 3.6 Hz).

Further comparisons between IWAIF model predictions and TCCT discrimination in the normal-hearing listeners are limited because of the number of conditions tested and because pure tone fdl's were not measured in our listeners. However, results from other studies support the idea that TCCT discrimination in normal-hearing listeners is consistent with predictions of the IWAIF model (Anantharaman *et al.*, 1993; Dai, 1993; Versfeld and Houtsma, 1995). In a group of normal-hearing listeners, Versfeld and Houtsma (1995) showed good correspondence between pure tone fdl's and ΔIWAIF values corresponding to threshold for changes in the level (ΔL) of TCCT signals at center frequencies of 1000 Hz and above.

The TCCT discrimination in listeners with hearing loss is not well described by the IWAIF model. The ΔIWAIF values based on the group psychometric functions of the hearing-impaired listeners vary substantially across conditions. (Similar variability was observed for ΔIWAIF values calculated for psychometric functions from individual hearing-impaired listeners.) While comparisons must be made cautiously due to large intersubject variability among hearing-impaired listeners, the TCCT discrimination deficits in our hearing-impaired listeners appear to be larger than pure tone frequency discrimination deficits reported for hearing-impaired listeners in the literature. Nelson and Freyman (1986) showed that pure tone fdl's in a group of listeners with mild to moderate hearing loss ranged from near normal to up to 15 times those of normal-hearing listeners. In a study of pure tone frequency discrimination in listeners with unilateral hearing loss, Moore and Glasberg (1986) reported that the average ratio of pure tone fdl's obtained in the impaired ears to those obtained in the normal ears was 2.3 at 1000 Hz and 3.1 at 2000 Hz. The ΔIWAIF values reported in Table IV for the hearing-impaired listeners are at least 20 times greater than the values reported for the normal-hearing listeners.

In summary, neither the excitation pattern model nor the IWAIF model can predict TCCT discrimination in both groups of listeners. Except for quantitative comparisons at 2000 Hz, TCCT discrimination in normal-hearing listeners is generally consistent with the IWAIF model and with the idea that the discrimination task in normal-hearing listeners is based on phase locking to the temporal fine structure of the TCCTs. A temporal processing deficit may be one possible explanation for reduced TCCT discrimination in the listeners with hearing loss: If hearing-impaired listeners have a reduced ability to use phase-locking information in TCCT discrimination, then they may rely more on place cues in carrying out this task.

D. Falling portion of psychometric function

Both group and individual psychometric functions for listeners with hearing loss do not show a decrease in discrimination ability at the largest values of Δf tested in this experiment (Figs. 2, 4, and 5). In contrast, group psychometric functions for listeners with normal hearing showed a worsening in performance at large Δf's. The falling portions of these group psychometric functions were well described by least-squares linear regression fits ($R^2=0.95$). The equations for these regression lines were used to determine the Δf corresponding to $d'=1$ on the falling portion of the psychometric functions. Defined here as the upper limit of discrimination, the values of Δf corresponding to $d'=1$ on the falling portion of the psychometric function are shown in Table V.

The upper limits of discrimination increase as f_c increases and as sensation level increases. Feth and O'Malley (1977) reported these same trends for the falling portions of the psychometric functions for discrimination they measured in a group of normal-hearing listeners. These results are consistent with the idea that spectral resolution plays a role in TCCT discrimination at larger Δf's. The upper limits of performance reported here and by Feth and O'Malley are substantially larger than empirical estimates of the critical band

TABLE V. Estimates of the upper limit of discrimination corresponding to a $d'=1$ on falling portion of group psychometric functions for normal-hearing listeners.

Condition	Upper limit of discrimination (Hz)
$f_c=1$ kHz, $\Delta L=3$ dB	1063
$f_c=1$ kHz, $\Delta L=1$ dB	493
$f_c=2$ kHz, $\Delta L=3$ dB	1359
$f_c=2$ kHz, $\Delta L=1$ dB	618
$f_c=4$ kHz, $\Delta L=3$ dB	1992
$f_c=1$ kHz, $\Delta L=1$ dB, 25 dB SL	290
$f_c=2$ kHz, $\Delta L=3$ dB, 25 dB SL	1028

(Zwicker *et al.*, 1957) or of the equivalent rectangular bandwidth (Moore and Glasberg, 1983). Discrimination ability presumably breaks down when the pure tone components of a TCCT are separated by an amount that is greater than the frequency separation required to hear the components as separate tones. Presumably, this breakdown occurs when there is insufficient overlap between the excitation patterns of the two components of a TCCT signal (Moore, 1998; Feth and O'Malley, 1977; Green, 1965).

Whereas group psychometric functions for listeners with normal hearing showed a worsening in performance at large Δf's, several listeners with normal hearing demonstrated excellent discrimination at the largest values of Δf tested in this study. The absence of a falling portion of the psychometric function in some normal-hearing listeners was most evident for $\Delta L=3$ dB. For example, Fig. 8 shows discrimination performance of listeners N2 and N3 for the falling portion of the psychometric function at $f_c=1000$ Hz, $\Delta L=1$ dB and at $f_c=1000$ Hz, $\Delta L=3$ dB. Listener N3 shows decreased performance at large Δf's in both conditions. In contrast, listener N2 shows improved performance at $f_c=1000$ Hz, $\Delta L=3$ dB, $\Delta f=1000$ Hz. This increased performance in $\Delta L=3$ dB conditions suggests that some listeners might utilize intensity cues in TCCT discrimination at Δf greater than 283 Hz.

Level cues might also affect the TCCT discrimination in listeners with hearing loss at large Δf's. Alternatively, listeners with hearing loss with decreased frequency selectivity might show the inverted U-shaped function if the range of Δf's was increased. To investigate this possibility, we tested TCCT discrimination at $\Delta f=1414$ Hz and $\Delta f=2000$ Hz in listeners I3, I4, and I9 at $f_c=1000$ Hz and in listeners I2, I7, I11, and I12 at $f_c=2000$ Hz. While most of these listeners did not show a change in performance at these very large values of Δf, listener I7 showed a substantial change in performance as the range of Δf's increased (Fig. 9).

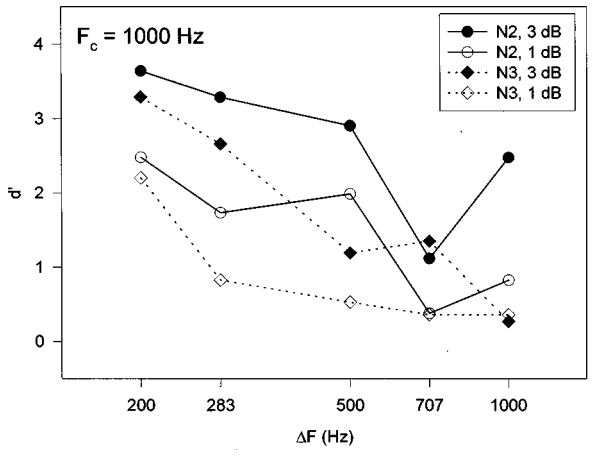


FIG. 8. Discrimination shown for listeners N2 and N3 for the falling portion of the psychometric functions at $f_c=1000$ Hz, $\Delta L=1$ dB and at $f_c=1000$ Hz, $\Delta L=3$ dB.

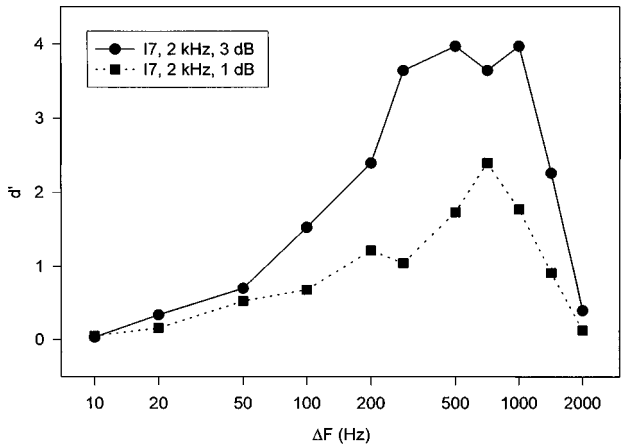


FIG. 9. Psychometric functions for listener 17 at $f_c=2000$ Hz, $\Delta L=1$ dB and at $f_c=2000$ Hz, $\Delta L=3$ dB for Δf's ranging from 10 to 2000 Hz.

Because in this study overall level was not randomized between TCCT signals presented on a single trial, listeners might focus on changes in the level of a single pure tone component present in both intervals of a trial. Thus, it is not possible to separate out the role of spectral resolution and the potentially confounding effects of level cues in describing TCCT discrimination at large values of Δf in the listeners tested in this study.

E. Possible age effects

Age might also contribute to the reduced overall performance of listeners with hearing loss on the TCCT discrimination task. Several investigators have suggested that the effects of increased age, independent of the effects of hearing loss, contribute to reduced frequency selectivity and temporal processing (e.g., Patterson *et al.*, 1982; Gordon-Salant and Fitzgibbons, 1993). Other researchers have reported that age per se does not appear to play a major role in reduced spectral and temporal processing (e.g., Lutman *et al.*, 1991; Moore *et al.*, 1992; Takahashi and Bacon, 1992). In this study, the listeners with hearing loss ranged from 36–70 years, whereas the listeners with normal hearing were all under 35 years of age. Different results might have been obtained if the listeners with normal hearing and the listeners with hearing loss were more closely matched in age. However, no systematic trend between age and performance is evident with the group of listeners with hearing loss. For example, in the $f_c=1000$ Hz, $\Delta L=3$ dB condition, the two listeners with the worst performance in the hearing-impaired group represented the oldest (I3, age 69) and youngest (I4, age 36) listeners. Similarly, at $f_c=1000$ Hz, the two young-

est listeners represented the best (I5, age 38) and worst (I4, age 36) performance within the group of listeners with hearing loss.

III. SUMMARY AND CONCLUSIONS

(1) Average TCCT discrimination for the group of listeners with normal hearing is described by inverted U-shaped psychometric functions in which discrimination improves as Δf increases, is (nearly) perfect for a range of Δf's, and then decreases again as Δf increases. In contrast, group psychometric functions for listeners with hearing loss are shifted to the right such that above chance performance occurs at larger values of Δf than in listeners with normal hearing. Furthermore, group psychometric functions for listeners with hearing loss do not show a decrease in performance at the largest values of Δf included in this study.

(2) Difference limens for TCCT discrimination were significantly larger for listeners with hearing loss than for listeners with normal hearing. This decreased discrimination ability is evident when listeners with hearing loss are compared to listeners with normal hearing both at equal SPLs and at equal sensation levels. Furthermore, in both listeners with normal hearing and in listeners with hearing loss, TCCT discrimination is significantly worse at 4000 Hz than at lower center frequencies and is worse for $\Delta L = 1$ dB than for $\Delta L = 3$ dB.

(3) The TCCT discrimination in listeners with normal hearing is inconsistent with an excitation pattern model. Data from normal-hearing listeners are in qualitative agreement with a temporal model of TCCT discrimination and show quantitative agreement with the IWAIF model at $f_c = 1000$ Hz but not at $f_c = 2000$ Hz.

(4) The TCCT discrimination in listeners with hearing loss is consistent with an excitation pattern model and is inconsistent with the IWAIF model. Listeners with hearing loss may have deficits in the temporal encoding of average instantaneous frequencies and so may rely more on place cues in TCCT discrimination.

(5) Falling portions of the psychometric functions at large Δf's suggest that the upper limit of performance of TCCT discrimination is related to spectral resolution. However, more research is needed to separate out the role of spectral resolution and the role of level cues in describing TCCT discrimination in normal-hearing and hearing-impaired listeners at large values of Δf.

ACKNOWLEDGMENTS

This research was supported by a grant from the Deafness Research Foundation. Portions of these results were reported at the Convention of the Acoustical Society of America, Indianapolis, Indiana, May 1996. Portions of the results were also part of a thesis submitted by the second author to the University of Colorado in partial fulfillment of the Master of Arts degree. The authors extend thanks to Tom Eberhard for programming and technical assistance and to Brian Moore for sharing his excitation pattern software. Lew Harvey, Brian Moore, and an anonymous reviewer also provided helpful comments and suggestions.

APPENDIX: EXCITATION PATTERN CALCULATIONS

This appendix provides a summary of the calculations used in determining excitation patterns with the model developed by Moore and Glasberg (1997). The reader is referred to Moore and Glasberg (1997, 1983) and Glasberg and Moore (1990) for further details regarding the calculation of auditory filter shapes and of excitation patterns. In the Moore and Glasberg (1997) model, the excitation pattern for each TCCT is based on the total output of auditory filters across a wide range of center frequencies. The total output of each auditory filter is determined by first calculating the filter output for each pure tone component in the TCCT stimulus (in terms of intensity) and then summing the output intensities for each of the components.

The output of each filter to the pure tone components in the TCCT stimuli is determined by the shapes of the auditory filters. If the center frequency of the filter is below the frequency of the pure tone component, then the upper branch (p_u) of the auditory filter determines the output of the filter to that tone. If the center frequency of the filter is above the frequency of the pure tone component, then the lower branch (p_l) of the auditory filter determines the output of the filter to that tone. Moore and Glasberg (1997) assumed that an auditory filter with a center frequency of F_c is described by two rounded exponential intensity-weighting functions (Patterson *et al.*, 1982) according to the following equations:

$$W(g) = (1 + p_l g) * \exp(-p_l g), \tag{A1}$$

$$W(g) = (1 + p_u g) * \exp(-p_u g), \tag{A2}$$

where p_u and p_l are parameters determining the sharpness of the upper and lower branches of the filter, respectively, and g is the deviation from F_c divided by F_c.

The parameters p_u and p_l for normally hearing listeners and for an input level of 51 dB per ERB are calculated by the equation

$$P = 4.0 * F_c / \mathrm{ERB}, \tag{A3}$$

where ERB is the equivalent rectangular bandwidth of the filter for that input level. The ERB of filters in normal-hearing listeners for an input level of 51 dB/ERB is defined by the equation

$$\mathrm{ERB} = 24.7(4.37F_c + 1). \tag{A4}$$

In the presence of cochlear-based hearing loss (quantified here in terms $\mathrm{HL_{bhc}}$ which is the amount of hearing loss due to OHC damage), the ERBs for filters with F_c of 1000 Hz and greater are assumed to be increased by a broadening factor, B:

$$B = 10^{0.01757(HL_{\mathrm{ohc}} - 22)}. \tag{A5}$$

Because cochlear hearing loss does not cause as much broadening of auditory filters centered at lower frequencies, the equation defining B is modified in the following way for filters centered below 1000 Hz:

$$B = 10^{[0.01757(HL_{ohc}-22)(1-(F_c-1)^2)/3.09]}. \quad \text{(A6)}$$

The steepness of the lower branch of the auditory filter decreases with increasing level in listeners with normal hearing. Moore and Glasberg include this effect of increasing level on p_l at a particular center frequency F_c through the following equation:

$$p_l(X) = p_{l(51)} - 0.35(p_{l(51)}/p_{l(51,1k)})(X-51), \quad \text{(A7)}$$

where X is the effective input level (in dB/ERB), $p_l(51)$ is the value of p_l at F_c for an input level of 51 dB/ERB and $p_l(51,1k)$ is the value of p_l at 1000 Hz for an effective input level of 51 dB/ERB. Because changes in the auditory filter shape with level are not as pronounced with cochlear hearing loss, the slope (0.35) in Eq. (A7) above becomes a function of the broadening factor B, in determining excitation patterns in impaired auditory systems:

$$\text{slope} = 0.35 - [0.35(B-1)/3]. \quad \text{(A8)}$$

[1]Following Macmillan and Creelman (1991), perfect discrimination scores (proportion correct=1.0) were converted to $1-1/2N$, where N is the total number of trials. In this study, $N=200$ trials, so proportion correct of 1.0 was converted to 0.9975, which corresponds to a $d'=4.0$.

[2]Equivalent rectangular bandwidth (ERB) number has been defined by Glasberg and Moore (1990) as a frequency scale in which number of ERBs=21.4*LOG$_{10}$((4.37*F)+1), where F is in kHz.

ANSI (**1989**). ANSI S3.6-1989, ''Specifications for audiometers'' (American National Standards Institute, New York).

Anantharaman, J. N., Krishnamurthy, A. K., and Feth, L. L. (**1993**). ''Intensity-weighted average of instantaneous frequency as a model for frequency discrimination,'' J. Acoust. Soc. Am. **94**, 723–729.

Arehart, K. H. (**1994**). ''Effects of harmonic content on complex-tone fundamental-frequency discrimination in hearing-impaired listeners,'' J. Acoust. Soc. Am. **95**, 3574–3585.

Burns, E. M., and Feth, L. L. (**1983**). ''Pitch of sinusoids and complex tones above 10 kHz,'' in *Hearing-Physiological Bases and Psychophysics*, edited by R. Klinke and R. Hartmann (Springer-Verlag, Berlin).

Carney, A. E., and Nelson, D. A. (**1983**). ''An analysis of psychophysical tuning curves in normal and impaired ears,'' J. Acoust. Soc. Am. **73**, 268–278.

Dai, H. (**1993**). ''On the pitch of two-tone complexes,'' J. Acoust. Soc. Am. **94**, 730–734.

Feth, L. L. (**1974**). ''Frequency discrimination of complex periodic tones,'' Percept. Psychophys. **15**, 375–378.

Feth, L. L., and O'Malley, H. (**1977**). ''Two-tone auditory spectral resolution,'' J. Acoust. Soc. Am. **62**, 940–947.

Feth, L. L., O'Malley, H., and Ramsey, Jr., J. (**1982**). ''Pitch of unresolved, two-component complex tones,'' J. Acoust. Soc. Am. **72**, 1403–1412.

Florentine, M., and Buus, S. (**1981**). ''An excitation-pattern model for intensity discrimination,'' J. Acoust. Soc. Am. **70**, 1646–1654.

Freyman, R. L., and Nelson, D. A. (**1991**). ''Frequency discrimination as a function of signal frequency and level in normal-hearing and hearing-impaired listeners,'' J. Speech Hear. Res. **34**, 1371–1386.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hearing Res. **47**, 103–138.

Gordon-Salant, S., and Fitzgibbons, P. J. (**1993**). ''Temporal factors and speech recognition performance in young and elderly listeners,'' J. Speech Hear. Res. **36**, 1276–1285.

Grant, K. W. (**1987**). ''Frequency modulation detection by normally hearing and profoundly hearing-impaired listeners,'' J. Speech Hear. Res. **30**, 558–563.

Green, D. M. (**1965**). ''Masking with two tones,'' J. Acoust. Soc. Am. **37**, 802–813.

Hacker, M. J., and Ratcliff, R. (**1979**). ''A revised table of d' for m-alternative forced choice,'' Percept. Psychophys. **26**, 168–170.

Harrison, R. V., and Evans, E. F. (**1979**). ''Some aspects of temporal coding by single cochlear fibers from regions of cochlear hair cell degeneration in the guinea pig,'' Arch. Otorhinolaryngol. **224**, 71–78.

Lutman, M. E., Gatehouse, S., and Worthington, A. G. (**1991**). ''Frequency resolution as a function of hearing threshold level and age,'' J. Acoust. Soc. Am. **89**, 320–328.

Macmillan, N. A., and Creelman, C. D. (**1991**). *Detection Theory: A User's Guide* (Cambridge U.P., Cambridge).

Moore, B. C. J. (**1995**). *Perceptual Consequences of Cochlear Damage* (Oxford Medical, New York).

Moore, B. C. J. (**1998**). Personal communication.

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. (**1986**). ''The relationship between frequency selectivity and frequency discrimination for subjects with unilateral and bilateral cochlear impairments,'' in *Auditory Frequency Selectivity*, edited by B. C. J. Moore and R. D. Patterson (Plenum, New York).

Moore, B. C. J., and Glasberg, B. R. (**1997**). ''A model of loudness perception applied to cochlear hearing loss,'' Aud. Neurosci. **3**, 289–311.

Moore, B. C. J., and Peters, R. W. (**1992**). ''Pitch discrimination and phase sensitivity in young and elderly subjects and its relationship to frequency selectivity,'' J. Acoust. Soc. Am. **91**, 2881–2893.

Moore, B. C. J., and Sek, A. (**1994**). ''Effects of carrier frequency and background noise on the detection of mixed modulation,'' J. Acoust. Soc. Am. **96**, 741–751.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1992**). ''Detection of temporal gaps in sinusoids by elderly subjects with and without hearing loss,'' J. Acoust. Soc. Am. **92**, 1923–1932.

Nelson, D. A., and Freyman, R. L. (**1986**). ''Psychometric functions for frequency discrimination from listeners with sensorineural hearing loss,'' J. Acoust. Soc. Am. **79**, 799–805.

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.

Plomp, R. (**1964**). ''The ear as a frequency analyzer,'' J. Acoust. Soc. Am. **36**, 1628–1636.

Plomp, R., and Mimpen, A. M. (**1968**). ''The ear as a frequency analyzer II,'' J. Acoust. Soc. Am. **43**, 764–767.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). ''Phase locked response to low frequency tones in single auditory nerve fibers of the squirrel monkey,'' J. Neurophysiol. **30**, 769–793.

Simon, H. J., and Yund, E. W. (**1993**). ''Frequency discrimination in listeners with sensorineural hearing loss,'' Ear Hear. **14**, 190–201.

Stover, L. J., and Feth, L. L. (**1983**). ''Pitch of narrow-band signals,'' J. Acoust. Soc. Am. **73**, 1701–1707.

Takahashi, G. A., and Bacon, S. P. (**1992**). ''Modulation detection, modulation masking, and speech understanding in noise in the elderly,'' J. Speech Hear. Res. **35**, 1410–1421.

Turner, C. W., and Nelson, D. A. (**1982**). ''Frequency discrimination in regions of normal and impaired sensitivity,'' J. Speech Hear. Res. **25**, 34–41.

Tyler, R. S., Fernandes, M., and Wood, E. J. (**1982**). ''Masking of pure tones by broad-band noise in cochlear pathology,'' J. Speech Hear. Res. **25**, 117–124.

Versfeld, N. J., and Houtsma, A. J. M. (**1995**). ''Discrimination of changes in the spectral shape of two-tone complexes,'' J. Acoust. Soc. Am. **98**, 807–816.

Wakefield, G. H., and Nelson, D. A. (**1985**). ''Extension of a temporal model of frequency discrimination: Intensity effects in normal and hearing-impaired listeners,'' J. Acoust. Soc. Am. **77**, 613–619.

Wier, C. C., Jesteadt, W., and Green, S. M. (**1977**). ''Frequency discrimination as a function of frequency and sensation level,'' J. Acoust. Soc. Am. **61**, 178–184.

Woolf, N. K., Ryan, A. F., and Bone, R. C. (**1981**). ''Neural phase locking in the absence of cochlear outer hair cells,'' Hearing Res. **4**, 335–346.

Zurek, P. M., and Formby, C. (**1981**). ''Frequency-discrimination ability of hearing-impaired listeners,'' J. Speech Hear. Res. **46**, 108–112.

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., Flottorp, G., and Stevens, S. S. (**1957**). ''Critical band width in loudness summation,'' J. Acoust. Soc. Am. **29**, 548–557.

Temporal integration and multiple looks, revisited: Weights as a function of time[a)]

Søren Buus
Communication and Digital Signal Processing Center, Department of Electrical and Computer Engineering (409 DA), Northeastern University, Boston, Massachusetts 02115

(Received 28 April 1997; revised 7 January 1999; accepted 14 January 1999)

This study tests the hypothesis that temporal integration for detection of tone bursts with various durations can be explained by optimally combining multiple looks of brief signal segments whose contribution to detection increases over time. Detectability was measured for signals consisting of six consecutive 25-ms, 1-kHz tone pulses presented in a 50-Hz-wide masker or in maskers consisting of seven 50-Hz-wide noises, one critical band apart, with either coherent or incoherent envelopes. The level of each signal pulse varied randomly around masked threshold according to a Gaussian distribution. The slopes of conditional psychometric functions—plotted in terms of d'^2 as a function of the squared signal-pulse intensity for pulses in a particular temporal position—yielded estimates of the contribution to detection provided by each pulse. Results for three normal listeners showed a small, but significant, effect of the temporal location of the pulse. Multiple-looks predictions of temporal-integration functions based on the measured weights and on measured psychometric functions were compared to measured temporal-integration functions. For the single-band and incoherent maskers, the predicted temporal-integration slopes were reasonably consistent with those measured, but for the coherent masker it was not. Whereas no current theory can explain the very steep temporal-integration functions obtained in the coherent masker, the present results are not inconsistent with the multiple-looks hypothesis as an explanation for the decrease in threshold with increasing duration of signals presented in random maskers. © *1999 Acoustical Society of America.* [S0001-4966(99)04404-5]

PACS numbers: 43.66.Mk, 43.66.Dc, 43.66.Ba [WJ]

INTRODUCTION

This paper tests the hypothesis that the well-known decrease in detection thresholds with increasing signal duration (e.g., Hughes, 1946; Plomp and Bouman, 1959; for review, see Florentine, Fastl, and Buus, 1988) reflects that the listener's decision is based on an optimal combination of multiple brief observations (multiple looks) rather than on the output of an integrator with a relatively long time constant.

Viemeister and Wakefield (1991) proposed a multiple-looks hypothesis to explain why the threshold for two brief pulses remained constant at −1.5 dB (*re*: threshold for a single pulse) when the pulses were more than about 5 ms apart. This finding has recently been replicated and extended by measurements for brief tone bursts (Krumbholz *et al.*, 1996; Wiegrebe *et al.*, 1996). These results are inconsistent with the common assumption that detection is based on the output of an integrator with a time constant of 100 to 200 ms (e.g., Green *et al.*, 1957; Plomp and Bouman, 1959; Buus, 1997). Whereas such an integrator easily accounts for the fact that thresholds decrease with increasing duration, at least up to a few hundred ms, it predicts that the detection threshold for two separate pulses should vary with their temporal separation well beyond the 5-ms limit observed by Viemeister and Wakefield (1991). The exact predictions depend on the form of the integrator. Plomp and Bouman's (1959) exponential-window integrator predicts that the threshold for two pulses separated by less than 25 ms should remain between −3 and −2.75 dB (*re*: threshold for a single pulse), reach −1.5 dB only as the separation increases to 175 ms, and approach 0 dB for separations exceeding 500 ms. Green *et al.*'s (1957; see also Buus, 1997) rectangular-window integrator with multiple observations predicts that the threshold for two pulses should remain at −3 dB for separations up to the 100-ms integration time, beyond which it should increase to −1.5 dB.

Viemeister and Wakefield (1991) also found that the threshold for two pulses separated by 100 ms was affected little by 6-dB increments or decrements in an intervening noise. This finding is consistent with detection being based on an optimal combination of multiple brief observations and inconsistent with long-term integration. According to Plomp and Bouman's (1959) model, decreasing or increasing the level of the intervening noise should affect the thresholds for two pulses 100 ms apart, because detection depends on the stimulus intensity preceding the end of the second pulse and variations in the masker intensity at times well within the integration window affect the output of the integrator.[1] On the other hand, an optimal combination of multiple, independent, brief observations excludes the intervening noise from the decision. Thus, changing the level of the noise should have no effect.

Viemeister and Wakefield (1991) also showed that the multiple-looks theory might account for temporal integra-

[a)]Parts of this paper were presented at the 21st Midwinter Research Meeting of the Association for Research in Otolaryngology, St. Petersburg Beach, FL, February 1998 [Abs. 21st Midwinter Res. Mtng. Assoc. Res. Otolaryngol. 202 (1998)].

tion; that is, the decrease in thresholds with increasing duration of relatively long sounds. If all temporal segments (looks) contribute equally to detection, the theory predicts that the threshold should decrease 5 dB per decade of signal duration, provided d' is proportional to signal intensity as is indicated by data for both short and long tones presented in the quiet or in broadband noise (e.g., Buus *et al.*, 1986; Carlyon *et al.*, 1990). The predicted threshold decrease of 5 dB per decade is smaller than the slope of about -10 dB/decade that is generally obtained for durations between the reciprocal of the critical bandwidth and about 100 ms, but matches the slope of about -5 dB/decade obtained for durations longer than 100 ms (e.g., Green *et al.*, 1957; Moore, 1989). However, if it is assumed that the contribution to detection by each look (i.e., the ''weight'') increases over the first 200 ms of the stimulus, the multiple-looks theory can predict the data.

Gilkey (1987) derived weights as a function of time by optimizing the parameters of a detection model to yield the best possible prediction of individual listeners' responses to tones presented in 25 exemplars of frozen noise. The results obtained with two signal levels were generally consistent within listeners, but considerable differences were apparent among listeners. Two listeners showed little effect of time within the 100-ms signal, one listener showed a markedly higher weight in the middle than in the beginning and end of the signal, and one listener showed weights that increased over the first 80 ms of the signal. Moreover, the temporal variation of weights accounted for only 6% of the variance in Gilkey's (1987) data. The large interlistener differences and the weak predictive power of the weights make them difficult to interpret. Therefore, this paper aims to use a different method for measuring how the contribution to detection by distinct temporal segments of a multisegment signal varies over the duration of the signal. To allow a complete evaluation of the multiple-looks theory as an explanation for the decrease in thresholds with increasing duration of a signal, thresholds as a function of duration and psychometric functions will also be measured.

The primary measurements will be based on a variation of a procedure recently developed by Buus, Zhang, and Florentine (1996). This procedure is similar to Berg's (1989) conditional-on-single-stimulus (COSS) procedure, except that perturbations are imposed on the signal and thus occur only in the signal interval. If the levels of distinct temporal segments of a multisegment signal are varied randomly and independently, the weight assigned to each temporal segment can be estimated by measuring the sensitivity, d', as a function of the perturbation for a given segment. Further details of this method and its application to the multiple-looks theory are given below.

I. DERIVATION OF DETECTION CONTRIBUTION ACCORDING TO THE MULTIPLE-LOOKS THEORY

This section considers how the detection contribution by each temporal segment of a signal can be estimated from measurements in which the levels of individual segments of the signal vary randomly and independently. For example, the signal may consist of a sequence of tone pulses, whose levels are chosen randomly. According to Viemeister and Wakefield's (1991) multiple-looks theory, each temporal segment may make a different contribution to the overall detectability of a multisegment signal. If Eq. (2) in their paper is generalized to accommodate different signal intensities in each look, one finds that the overall sensitivity, d', should be related to the intensities, I_i, of the signal during each of the looks as follows:

$$d'^2 = \sum_{i=1}^{N} k_i^2 I_i^2, \tag{1}$$

where N is the number of looks applied to the signal, i is the number of the look (which will be equated to the pulse number in the present experiment),[2] and k_i is the detection contribution or weight for the ith segment of the signal.

To find the relative contribution to detection by the nth pulse, k_n, Eq. (1) can be rewritten as

$$d'^2 = k_n^2 I_n^2 + \sum_{i \neq n} k_i^2 I_i^2 = k_n^2 I_n^2 + C_n. \tag{2}$$

This shows that d'^2 plotted as a function of I_n^2 should be a straight line whose slope is proportional to the squared detection contribution for signal segment n, provided that k_n is independent of I_n. The intercept, C_n, is a random variable because I_i is random. However, that does not affect the estimated slope of the conditional psychometric function, because the intercept represents the square of the total contribution to the overall sensitivity provided by the signal segments that are not under consideration.

II. TEMPORAL INTEGRATION AND WEIGHTS AS A FUNCTION OF TIME

To test the hypothesis that the decrease in thresholds with increasing duration of a signal can be explained by a multiple-looks theory in which independent looks are combined with weights that vary with time since the onset of the signal, signal segments will be selected according to their temporal position within a multisegment signal. This allows conditional psychometric functions to be estimated for various times after the onset of the signal. The slopes of these functions will reveal if the contribution to detection by a signal segment varies with its temporal location within the signal.

A. Method

In a previous paper, Buus *et al.* (1996) investigated whether time-varying weights might explain comodulation masking release (CMR) by measuring weights as a function of short-term masker level in three masking conditions: a narrow-band masker presented alone (a single-band masker), together with six flanking bands with envelopes that were unrelated to the masker band and to one another (an incoherent masker), or together with six flanking bands that had envelopes identical to that of the masker band (a coherent masker). The weights were obtained by sorting signal segments according to the short-term masker level in a 20-ms window around the signal segment. The results showed that

the weights were nearly independent of the short-term masker level for the single-band and incoherent maskers, but decreased markedly with increasing short-term masker levels for the coherent masker. In other words, they revealed clearly different listening strategies for the incoherent and coherent maskers. For the incoherent masker, all parts of the signal were equally important for detection. For the coherent masker, the parts of the signal occurring during periods of low short-term masker energy contributed much more to detection than those occurring during periods of high short-term masker energy, which is consistent with ''listening in the valleys.'' These results show that conditional psychometric functions can be used to assess the listening strategy used for a particular task.[3]

Buus *et al.* (1996) sorted their data according to short-term energy of the masker because they were interested in determining how the contribution to detection (or weight) depended on the properties of the masker. If the same data are sorted according to the temporal position of the segments, the weight as a function of time may be obtained. Accordingly, the primary data reported in the present study result from a reanalysis of the responses obtained by Buus *et al.* (1996). To verify that temporal integration for the multisegment signal was not substantially different from that obtained for single tone pulses of various durations, an initial experiment measured threshold as a function of signal duration by varying the number of pulses in the signal. Because the method and most of the raw data are the same as those used by Buus *et al.* (1996), the following description is relatively brief and emphasizes aspects that differ from the previous paper.

1. Stimuli and apparatus

The signal was composed of 1-kHz tone pulses presented without pauses between them. Each pulse consisted of a 25-ms raised-cosine rise followed immediately by a 25-ms raised-cosine fall, which yielded a duration of 25 ms measured between the half-amplitude points of the envelope and an equivalent rectangular duration of 18.75 ms. This signal can also be considered a tone that is sinusoidally amplitude modulated with a modulation frequency of 20 Hz. In the main experiment, the signal always consisted of six pulses. To allow measurements of temporal integration, signals comprising one, two, four, and six pulses (modulation periods) were used in an initial experiment.

The masker was a 50-Hz-wide band of noise, arithmetically centered at 1 kHz, which was produced by addition of 25 sinusoids separated by 2 Hz. For each sinusoid, the amplitude was chosen at random from a Rayleigh distribution and the phase from a uniform distribution. In the incoherent and coherent masker conditions, six 50-Hz-wide bands centered at 0.57, 0.7, 0.84, 1.17, 1.37, and 1.6 kHz were added. Each of the coherent flanking bands was produced by reusing the amplitudes and phases for the sinusoids in the masker band. The six incoherent flanking bands were produced by choosing 150 new amplitudes and phases independently. The level of each noise band was set to 80 dB SPL. The masker and flanking bands had a duration of 500 ms, measured from the beginning of the 20-ms rise to the end of the 20-ms fall. Their onset occurred 150 ms before the onset of the signal. The apparatus was that used by Buus *et al.* (1996).

2. Procedure

Listeners' ability to detect the signal was measured in a two-interval, two-alternative, forced-choice paradigm. Each trial consisted of two 500-ms observation intervals marked by maskers and lights. The observation intervals were separated by a 500-ms interstimulus interval. The signal was presented in either the first or the second interval with equal *a priori* probability. The listener's task was to press a button corresponding to the interval that contained the signal. Immediately after the listener's response, the correct answer was indicated visually on the response terminal, after which the next trial followed with a short delay.

To examine the temporal-integration function for the multisegment signals, thresholds for signals composed of one, two, four, and six consecutive pulses of the 1-kHz tone were measured with an adaptive procedure. All tone pulses within a signal had the same level, which decreased after three consecutive correct responses and increased after one incorrect response. The step size was 5 dB initially, but was reduced to 2 dB after the second reversal. Reversals occurred whenever the signal level in successive trials changed from increasing to decreasing or vice versa. This procedure converged on the signal level necessary to produce 79.4% correct responses (Levitt, 1971). A single threshold measurement consisted of three brief, consecutive tracks. Each track terminated after five reversals and the threshold was calculated from the average signal level at the last two reversals for each of the three tracks. For each masker and signal condition, three threshold measurements (for a total of nine tracks) were obtained in random order.

Weights as functions of time were derived from Buus *et al.*'s (1996) measurements of detectability of the six-pulse signal in a modified method of constant stimuli. The level of each pulse in the signal varied randomly around an average signal level, which was chosen individually for each listener and condition to yield about 80% correct responses. The level of each pulse was chosen independently from a Gaussian distribution with a root-mean-square, rms, perturbation (standard deviation) of 3 dB in one condition and 6 dB in another. For each listener, masker type, and rms perturbation, 25 blocks of 200 trials were obtained in mixed order, except that all data with the 3-dB rms perturbation were collected before the data with the 6-dB rms perturbation. Because the level of each of the six tone bursts in a single signal varied independently, each trial yielded data for six combinations of temporal location (pulse number) and signal-pulse level. Thus, the 5000 trials obtained for each listener, masker type, and rms perturbation allowed 30 000 combinations of temporal location and signal-pulse level to be examined.

3. Listeners

One male (L1) and two female listeners (L2 and L3) with normal hearing participated in the experiment. They were university students and were 22–29 years old. All lis-

teners had audiometric thresholds less than 10 dB HL between 0.25 and 8 kHz and had no history of hearing difficulty in the test ear. The left ear was used for L1 and the right ear for L2 and L3.

4. Data analysis

To investigate possible differences in temporal integration among the three maskers, a three-way [listener×masker ×log(number of pulses)] analysis of covariance (ANCOVA) for repeated measures was performed on the average threshold differences between the six-pulse signals and the shorter-duration signals for each listener and masker (Data Desk 5.0.1, Data Description, Inc., Ithaca, NY, 1995). In addition, Scheffé *post hoc* tests for contrast (Data Desk 5.0.1, 1995) were performed as appropriate to explore the sources of significant effects and interactions. Differences are reported as significant when $p<0.05$.

The 5000 responses obtained for each listener, masker type, and rms perturbation for the randomly perturbed six-pulse signals were analyzed according to temporal location (or pulse number) of each signal segment. First, the responses were sorted into 60 categories comprising ten ranges of signal-pulse intensity for each of the six possible temporal locations. The ranges of signal-pulse intensity were chosen such that approximately equal numbers of responses fell into each range. Because the response was recorded once for each of the six tone pulses in the signal for a given trial, approximately 500 responses (5000 trials×6 pulses/trial divided among 60 categories) were tallied in each pulse-number/pulse-intensity category for each listener and condition. Then, the sensitivity, d', was calculated from the conditional response tallies (response given the signal interval; i.e., 1|1, 2|1, 1|2, and 2|2) obtained for each pulse-number/pulse-intensity category. Finally, d'^2 was plotted as a function of the squared average signal-pulse intensity within the category in accordance with Eq. (2). Separate functions were plotted for each listener, masker type, rms perturbation, and pulse number. A line was fitted to each conditional psychometric function using the method of least-squared error and the relative contribution to detection by the nth pulse, k_n, was derived from the slopes as

$$k_n=\frac{a_n}{\sum_{i=1}^{6} a_i}, \tag{3}$$

where a_i is the square root of the slope of the conditional psychometric function for the ith pulse of the six-pulse signal for a particular listener, masker type, and rms perturbation. This method of normalizing the weights yields a sum of relative contributions across the six pulses equal to unity, such that the time-averaged weight across the six pulses is 0.167 for a given masker and listener.

To assess the effects of pulse number and masker condition on the contribution to detection, a four-way (listener ×pulse number×masker condition×rms perturbation) analysis of variance (ANOVA) for repeated measures (Data Desk 5.0.1, 1995) was performed on the normalized detection contributions. The square of the standard error of each contribution, calculated from the standard error of the slopes estimated by the LINEST function (EXCEL 5.0a, Microsoft Corp., Redmond, WA, 1995), was used as an estimate of the variance for each datum. Again, Scheffé *post hoc* tests for contrast (Data Desk 5.0.1, 1995) were performed as appropriate to explore the sources of significant effects and interactions. Differences are reported as significant when $p<0.05$.

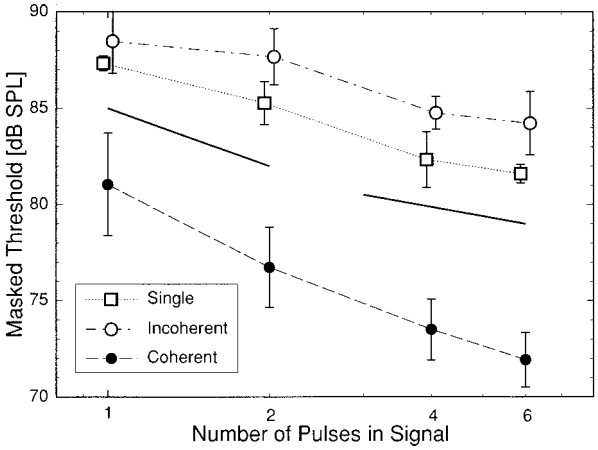


FIG. 1. Masked thresholds as a function of the number of consecutive 25-ms tone pulses in the signal. Separate functions are shown for the single-band (squares), incoherent (unfilled circles), and coherent maskers (filled circles). The error bars show plus and minus one standard deviation calculated across the data for the three individual listeners. Solid lines with slopes of −10 (left) and −5 dB/decade (right) are shown for comparison.

B. Results

1. Adaptive thresholds

Figure 1 shows the average masked thresholds for the three listeners. For each of the three maskers, the thresholds are plotted as a function of the number of pulses in the signal. As in other experiments on CMR (e.g., Hall *et al.*, 1984; for review, see Buus, 1997), the thresholds are highest in the incoherent masker and lowest in the coherent masker. The threshold difference, which is the CMR, increases from 7.4 dB for a single 25-ms pulse to 10.9 dB for two pulses, 11.3 dB for four pulses, and 12.3 dB for six pulses. The solid line segments show temporal-integration slopes of −10 (left) and −5 (right) dB/decade. As expected, the slopes of the temporal-integration functions tend to be steeper for small numbers of pulses than for larger numbers of pulses. Between four and six pulses, the slopes for single-band and incoherent maskers approach the −5 dB/decade expected for multiple observations of equally detectable pulses provided that d' is proportional to the signal intensity. For smaller pulse numbers, the slopes for the single-band and coherent maskers approach the −10 dB/decade predicted by an energy-detector model with long-term integration. The slopes of the best-fitting lines to the data for all pulse numbers are −6.0 dB/decade for the incoherent masker, −7.7 dB/decade for the single-band masker, and −11.7 dB/decade for the coherent masker.

The ANCOVA showed significant effects of masker type [$F(2,4)=8.15$, $p=0.039$] and log(number of pulses) [$F(1,2)=15\,500$, $p\leqslant 0.0001$], as well as a significant interaction of masker type and log(number of pulses) [$F(2,4)$

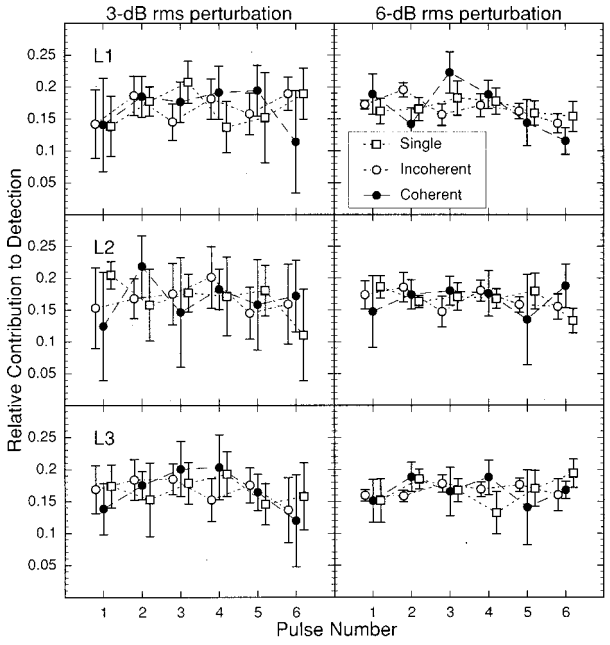


FIG. 2. Relative contributions to detection measured for each of the six pulses in the signal are shown as a function of its sequential location within the signal. Each row shows results for a different listener. The left column shows results for the 3-dB rms perturbation and the right column shows results for the 6-dB rms perturbation. Within each panel, separate functions are shown for the single-band (squares), incoherent (unfilled circles), and coherent maskers (filled circles). The error bars show the standard error of the contribution.

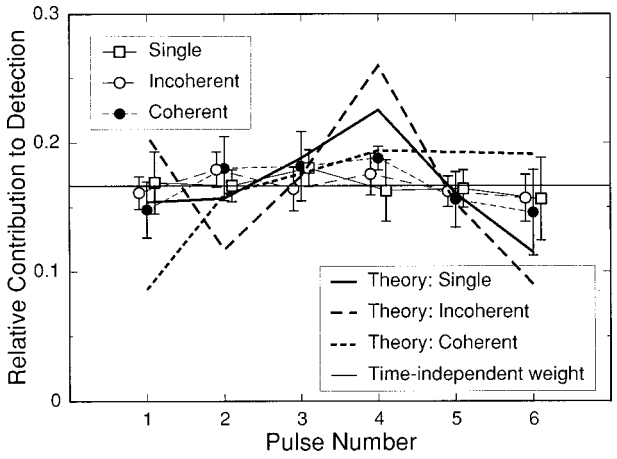


FIG. 3. Relative contributions to detection averaged across the three listeners and two rms perturbations. The average detection contributions are shown as functions of sequential location of the pulse. Separate functions are shown for the single-band (squares), incoherent (unfilled circles), and coherent maskers (filled circles). The error bars show plus and minus one standard deviation calculated across the six values for three listeners and two rms perturbations. The thick lines show the relative contributions necessary for the multiple-looks model to predict the temporal-integration functions in Fig. 1. Separate functions are shown for each masker condition. The thin solid line shows the time-averaged detection contribution across the six bursts, which is equal to 0.167. If the weights were completely independent of temporal location of the burst, all the data would lie on this line.

$=8.15$, $p=0.039$]. The effect of listener and all interactions with it were not significant. These results reflect that differences among the three listeners were relatively small, but the temporal-integration slopes varied among the three maskers. The Scheffé tests for contrast showed that the coherent masker yielded a significantly steeper slope than the incoherent and single-band maskers, whose slopes did not differ significantly.

2. Conditional psychometric functions

Figure 2 shows how each of the six signal pulses contributes to detection for each listener, masker, and rms perturbation. All three listeners show similar patterns of results. No systematic differences among the three maskers are apparent and the relative contribution to detection appears to be roughly the same for each of the six pulses.[4] The only notable trend is that the standard errors are considerably smaller for the 6- than for the 3-dB rms perturbation.

Figure 3 shows the contribution to detection as a function of pulse number averaged across the three listeners and two rms perturbations. As indicated in the top-left legend, separate functions are shown for the single-band, incoherent, and coherent maskers. The thin horizontal line shows the value of 0.167, which is the value that would be obtained if the weights were completely independent of time. The thick lines show the relative-contribution functions necessary for the multiple-looks theory to predict the temporal-integration functions in Fig. 1, as discussed later. The data show that the relative contribution to detection varies little across the six pulses in the signal. The data for the single-band and incoherent maskers show only small deviations from the thin line. The data for the coherent masker vary somewhat more, and the detection contributions tend to be highest for the second, third, and fourth pulses.

The ANOVA showed a significant effect of pulse number [$F(5,10)=15.77$, $p=0.0002$]. The Scheffé tests for contrast showed that the detection contributions of the second, third, and fourth pulses were higher than those of the first, fifth, and sixth pulses; all other differences were not significant. However, the deviation from equal detection contribution by each of the six bursts, although significant, is small. The ANOVA also showed a significant interaction between listener, pulse number, and masker type [$F(20,20)=2.68$, $p=0.016$], but the Scheffé tests for contrast showed no significant difference among any of the paired comparisons. Thus, no systematic differences among listeners are apparent.

C. Discussion

1. CMR as a function of duration

Figure 1 indicates that CMR depends on signal duration because the thresholds decrease more rapidly with duration in the coherent masker than in the incoherent and single-band maskers. Accordingly, the CMR is smaller at short than at long durations. The CMRs obtained in the present study are in good agreement with those obtained by Schooneveldt and Moore (1989) for 2000-Hz tone bursts with durations between 25 and 400 ms and a masker bandwidth of 1600 Hz, which is close to the six critical bands encompassed by the present maskers. On the other hand, McFadden (1986) found almost no effect of signal duration on CMR for 3715-Hz tone bursts with durations between 75 and 600 ms. In addition to the different signal frequencies used in these studies,

it should be noted that the number of bands and the envelope properties of the maskers also differed. It is not clear if these differences may explain, at least in part, why Schooneveldt and Moore (1989) and the present study found that signal duration affected the CMR, whereas McFadden (1986) found no effect. However, it seems that the different ranges of durations are the most likely cause of the discrepant findings. The results for all three studies agree that CMR is nearly independent of duration for durations exceeding 50–100 ms, and the two studies using shorter durations both found that the CMR decreased as the duration decreased below 50 ms.

2. Multiple looks and temporal integration

The most important finding in the present experiment is that only a small, although significant, effect of temporal location on the detection contribution by individual temporal segments of the signal was observed. This effect of temporal location is qualitatively similar to the effect of time (normalized to the total duration of the signal) on the average of the weights obtained by Gilkey (1987), but the present data vary much less over time and among listeners than do Gilkey's weights.

The present experiment yielded weights that were quite consistent both within and among listeners. Whereas such consistency usually is a desirable outcome for an experimental method, it could result if aspects of the method cause the listeners to change their listening strategy in some consistent manner. For example, it seems possible that modulating and perturbing the amplitude of the signal in the present experiment might affect the listening strategy, which could cause the measured weights to differ from those used for a steady tone of long duration. If the sinusoidal amplitude modulation changed the listening strategy, however, the temporal-integration functions obtained for our stimuli should also differ from those obtained for steady tones of various durations, but they did not. Thus, it seems unlikely that the modulation of the signal in the present experiment caused the weights to differ from those used for steady tones.

If the perturbation changed the listeners' strategy, one would expect the weights to differ between the 3- and 6-dB rms perturbations, but they did not. If the listeners depended on detecting a single random-position burst that happened to be perturbed to a high intensity, one would expect low- and moderate-intensity bursts not to affect listeners' performance. However, an analysis in which the pulse-number/pulse-intensity categories with high intensities were eliminated showed that d' increased with increasing intensity, even for bursts of low and moderate intensities. Thus, it appears that the perturbation of the signal did not affect the listeners' strategy to any appreciable extent.

Given that detection contributions show only a slight dependence on the temporal location of the signal segment, the multiple-looks theory predicts that thresholds should decrease only about 5 dB/decade, provided that d' is proportional to the signal intensity. This slope is less than those measured, at least for the single-band and coherent maskers. If the psychometric functions have similar shapes across the three masker conditions, the theory also predicts that temporal-integration functions obtained with the three maskers ought to be quite similar, given that the weight functions were similar. Whether these possible discrepancies between theory and data are important depends on the form of the psychometric functions. Therefore, the psychometric functions should be considered further.

It is well established that d' is proportional to signal intensity for signals presented in broadband maskers (e.g., Buus *et al.*, 1986; Carlyon *et al.*, 1990), but this may not be true for the maskers used in the present experiment. Moore *et al.* (1990) measured psychometric functions for 1-kHz tones presented in one and three 20-Hz-wide noise bands. The three bands had coherent envelopes. The results obtained with both maskers indicate that d' may not be proportional to the signal intensity. If d' is calculated from the percentages of correct responses (under the assumption of unbiased responses) and a line is fitted to $\log(d')$ plotted as a function of log(signal intensity), the best-fitting slopes average 0.62 for the narrow-band masker and 0.80 for the coherent masker. These slopes are shallower than the slope of unity that is obtained when d' is proportional to the signal intensity. Therefore, it appears that d' may grow more slowly than proportionally with intensity for signals presented in narrow-band maskers.

Moore *et al.* (1990) produced their maskers by multiplying a 10-Hz-wide low-pass noise by one or three sinusoids. Therefore, the maskers had envelopes corresponding to those of 10-Hz-wide Gaussian noise bands (yielding a rather large variance of short-term energy) and zero crossings that were regularly spaced in time. Either or both of these factors could be responsible for the somewhat shallow slopes of their psychometric functions. The present maskers are considerably less variable and do not have regular zero crossings. Therefore, it is unclear whether the present maskers might produce psychometric functions similar to those obtained by Moore *et al.* (1990) or similar to those obtained in broadband noises. The next experiment examines this question.

III. PSYCHOMETRIC FUNCTIONS

A. Method

1. Stimuli and apparatus

The maskers and signals were the same as in the previous experiments, except that minor differences in the maskers may have resulted from using different apparatus and signal-generation procedures. Psychometric functions were measured for signals consisting of one, two, four, and six consecutive 25-ms tone pulses presented in the single-band masker and, in one listener, for a four-pulse signal in the incoherent and coherent maskers. A PC-compatible computer with a signal processor (TDT AP2) generated the stimuli, sampled the listeners' responses, and executed the psychophysical procedure. Both the signal and the masker were generated on-line. The signals were generated in the time domain. The maskers were generated in the frequency domain. For each presentation, the real and imaginary parts of the spectral lines falling within the 50-Hz-wide passband(s) were set to random values chosen from a Gaussian distribution. For the coherent masker, the random values chosen for the first band were reused for the remaining six

bands. For the incoherent masker, random values were chosen independently for each of the seven passbands. The resulting complex spectrum was inverse-Fourier transformed to produce a 655.4-ms waveform, which was truncated and shaped with a 20-ms raised-cosine rise and fall to produce the 500-ms masker.

The signal and the masker were output through separate channels in a 16-bit D/A converter (TDT DD1) using a 50-kHz sample rate. The output of the signal channel was attenuated (TDT PA4), low-pass filtered (TDT FT5, $f_c = 20$ kHz, 135 dB/octave), and attenuated again (TDT PA4) before it was added (TDT SM3) to the masker and led to a headphone amplifier (TDT HB6), which fed one earphone of the Sony MDR-V6 headset. The output of the masker channel was attenuated (TDT PA4), low-pass filtered by the second channel of the TDT FT5, and led to the summation amplifier. For routine calibration, the output of the headphone amplifier was led to an A/D converter (TDT DD1), such that the computer could sample the waveform, calculate its spectrum and rms voltage, and display the results at the onset of each session.

2. Procedure

Psychometric functions were obtained by the method of constant stimuli. For each condition, ten signal levels, 1.5 dB/apart, were chosen to encompass the range from 55% to 95% correct responses. Each block of trials consisted of four warm-up trials (two each at the two highest signal levels) followed by 100 trials in which the signal level was chosen randomly (without replacement) on each trial to yield ten trials at each of the ten test levels. For each listener and condition, at least 12 blocks of trials were obtained. To ensure that the data for each listener and condition were relatively homogeneous, blocks that yielded an average percentage correct outside the 80%-confidence interval were eliminated from analysis. The confidence interval was calculated on the basis of the expected standard deviation for the mean of a binomial process with a mean equal to the average percentage correct obtained across all blocks of trials for the condition—defined as a combination of masker type and signal duration. In general, this procedure eliminated about 20% of the blocks. To ensure that each datum was based on at least 100 trials, extra blocks of trials were run if necessary until at least ten blocks were retained.

3. Listeners

Because the listeners for the previous experiment were no longer available, one female (L6) and two male (L4 and L5) listeners with normal hearing were recruited from the laboratory staff. They had no history of hearing difficulties and their audiometric thresholds were 15 dB HL or less between 0.25 and 8 kHz. Their ages ranged from 27 to 47 years and all were highly trained in psychoacoustic experiments. Listener L4 is the author. The right ear of L6 and the left ears of L4 and L5 were tested.

B. Results and discussion

Psychometric functions for signals consisting of one, two, four, and six pulses (modulation periods) presented in the single-band masker are shown in Fig. 4. Data are shown only for signal levels yielding performance significantly better than chance and significantly worse than perfect. Within this range, the logarithm of d' as a function of signal level reasonably can be summarized by sets of parallel lines as shown by the thick lines. The lines are fitted to the data by the method of least-squared error using a single value for the slope (i.e., the exponent when expressing d' as a power function of signal intensity) for each listener, but separate intercepts for each signal. The best-fitting slopes are 0.7 for L4, 0.72 for L5, and 0.84 for L6. The average slope is 0.76. Additional measurements in listener L4 using the four-pulse signal yielded psychometric functions with slopes of 0.72 in the incoherent masker and 0.70 in the coherent masker.

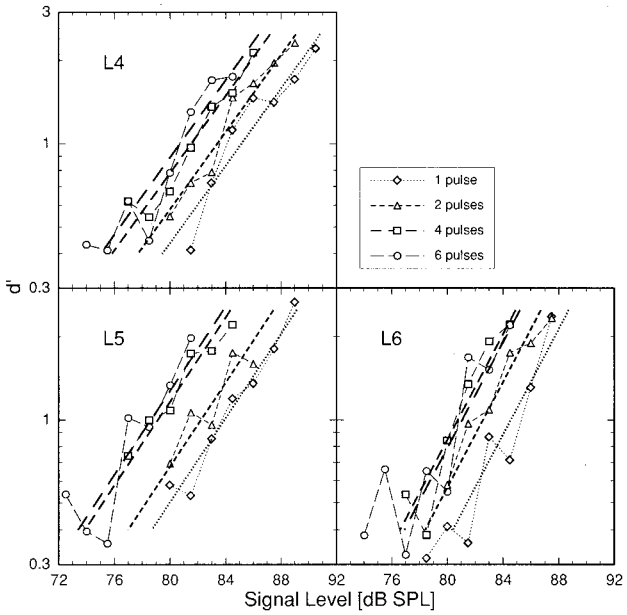


FIG. 4. Psychometric functions for detection of one, two, four, or six consecutive 25-ms tone pulses presented in the single-band masker. The sensitivity, d', is plotted on a logarithmic scale as a function of the signal level. The thick lines show the best-fitting parallel power functions to the data for each listener. The dotted lines are for one-pulse signals, the short-dashed lines for two-pulse signals, the medium-dashed lines for four-pulse signals, and the long-dashed lines for six-pulse signals. Each panel shows data for a single listener.

As discussed previously, Moore *et al.* (1990) obtained average slopes of 0.62 for their narrow-band masker and 0.80 for their comodulated masker. The slope for the narrow-band masker is somewhat shallower than that obtained in the present experiment, which may reflect that the masker bandwidth was considerably narrower in their experiment than in the present experiment. Their finding of steeper slopes for the coherent masker than for the narrow-band masker was not replicated by the single listener tested with both maskers in the present experiment, but this may reflect individual differences and, perhaps, variability in the data. Two of Moore *et al.*'s (1990) six listeners showed almost no difference in slopes between the two maskers. In any event, the slopes of the psychometric functions clearly are somewhat

less than unity for all of the maskers used in the present experiment.

IV. GENERAL DISCUSSION

The finding of psychometric-function slopes less than unity indicates that the slope of temporal-integration data in the same maskers should be somewhat steeper than −5 dB/decade if the weights are independent of time. Because the present data indicate that the psychometric-function slopes are approximately equal for all three maskers, the average slope of 0.76 for the single-band masker may be used as the psychometric-function slope for all the maskers. Given this slope and assuming multiple observations with time-independent weights, the expected temporal-integration slope is about −6.6 dB/decade for all three maskers. This slope is close to those obtained for temporal integration in the single-band and incoherent maskers, but shallower than that obtained in the coherent masker.

More precise predictions of the temporal-integration functions can be obtained by applying the average detection contributions shown in Fig. 3 to Eq. (1) with $d'=1$ and a psychometric-function slope of 0.76. As discussed above, the temporal-integration functions may differ among the three maskers, at least for short signal durations. Therefore, separate predictions were generated for each of the three maskers, although the detection contributions do not differ significantly among them. The predicted slopes were −6.5 dB/decade for the single-band masker, −6.8 dB/decade for the incoherent masker, and −7.6 dB/decade for the coherent masker, whereas the measured slopes were −7.7 dB/decade for the single-band masker, −6.0 dB/decade for the incoherent masker, and −11.7 dB/decade for the coherent masker. Statistical analysis showed that the predicted slope was significantly shallower than the measured slope for the coherent masker ($t=4.49$, $df=4$, $p=0.01$). The differences were not significant for the single-band and incoherent maskers. Thus, the multiple-looks theory may be considered approximately consistent with the data for the single-band and incoherent maskers, but it is inconsistent with the data for the coherent masker.

Whereas data and predictions for the slopes of the temporal-integration functions agree reasonably for the single-band and incoherent maskers, the agreement between the measured and the predicted weights is less clear. As shown by the thick lines in Fig. 3, the weights predicted on the basis of the thresholds obtained in experiment 1 and the average psychometric-function slope of 0.76 vary much more over time than those measured. In particular, the predicted weights in the single-band and incoherent maskers are considerably higher than those measured for the fourth pulse, and considerably lower for the sixth pulse. Likewise, the predicted weight for the first pulse of the signal in the coherent masker is much lower than that measured. The deviations between the predicted and measured weights are as large or larger for the single-band and incoherent maskers as for the coherent masker, even if the predicted average slopes for the former maskers were close to those obtained. This may indicate that the overall slope is an insensitive indicator of agreement between theory and measurement, or that the predicted weights are sensitive to perturbations in the measured thresholds.

To determine how much random perturbations in thresholds affect the weights estimated from the threshold functions, a simulation was performed. To determine reasonable ranges for perturbing thresholds in the simulation, the amounts of temporal integration were calculated by subtracting the threshold for six pulses from all the thresholds for each listener and masker. The normalized temporal-integration functions were then averaged across listeners and across the single-band and incoherent maskers to obtain the most stable estimate of the amount of temporal integration for conditions that did not differ significantly. Next, the weights were calculated for 5000 temporal-integration functions obtained by choosing randomly from Gaussian distributions with means and standard deviations equal to the averages and standard errors of the normalized data. The standard errors for the amounts of temporal integration for signals consisting of one, two, or four pulses were between 0.3 and 0.5 dB (the normalization caused it to be 0 dB for six pulses). Finally, the means and standard deviations of the 5000 weights estimated for each of the six pulses were calculated. The results showed that the mean weights were close to the mean of the two functions shown in Fig. 3 and the standard deviations ranged from about 0.02 for the weight on the first pulse to just over 0.05 for the weight on the fourth and sixth pulses. These amounts of variability indicate that the weights estimated from the masked thresholds for the single-band and incoherent maskers do not differ significantly from those measured, despite the poor fit between the estimated and measured weights shown in Fig. 3.

In conclusion, the present results for the single-band and incoherent maskers may be considered reasonably consistent with Viemeister and Wakefield's (1991) multiple-looks hypothesis, whereas the data for the coherent masker are not. However, no current theory of temporal integration can account for the very rapid decrease of the thresholds observed for the coherent masker and, indeed, the detection mechanism used in this masker may differ from that employed for truly random maskers. Therefore, the import of the discrepancy between theory and data for the coherent maskers is unclear and one may consider the present results not inconsistent with the multiple-looks hypothesis as an explanation of temporal integration for detection, at least for signals presented in random maskers.

V. SUMMARY

The present experiment measured conditional psychometric functions for detection of six-pulse signals to determine how the contribution of a short temporal segment of a long signal varies over the duration of the signal. Within the context of Viemeister and Wakefield's (1991) multiple-looks model, the square roots of the slopes of linear functions fitted to the data for d'^2 as a function of the squared signal intensity, I^2, provide estimates of the weights or relative contribution to detection for each segment of the signal. The weights measured in this manner were compared to weights

predicted from measurements of temporal integration and psychometric functions. The major findings are:

(1) For maskers consisting of seven 50-Hz-wide bands of Gaussian noise spaced one critical band apart, the temporal-integration function is shallower when the envelopes are incoherent than when they are coherent, especially at short signal durations. Thus, the CMR is larger for long than for short signals.
(2) The detection contribution of a short signal segment is approximately independent of its temporal location within a multisegment signal. This holds whether the masker is coherent, incoherent, or consists of only a single 50-Hz-wide band of noise.
(3) Psychometric functions indicate that d' grows more slowly than proportionally with signal intensity for the maskers used in the present experiment. The growth of d' can be approximated by a power function of signal intensity with an exponent of 0.76.
(4) The slopes of temporal-integration functions obtained in the single-band and incoherent maskers do not differ significantly from those predicted by the multiple-looks model with the measured weights and the measured psychometric function. However, the temporal-integration function in the coherent masker is significantly steeper than that predicted by the multiple-looks model.
(5) The present results are not inconsistent with the multiple-looks hypothesis as an explanation for the decrease in threshold with increasing duration of signals presented in random maskers. The results for the coherent masker do not support the multiple-looks hypothesis, but no current theory readily accounts for temporal integration in a coherent masker.

ACKNOWLEDGMENTS

Leiji Zhang and Debra Ingenito helped with data collection and analysis. Andrew Oxenham and Michelle Hicks helped collect the data in Fig. 4. Mary Florentine, Hannes Müsch, and two anonymous reviewers gave helpful comments on an earlier version of this paper. This research was supported by NIH/NIDCD grant number R01 DC00187.

[1]Predictions for Green *et al.*'s (1957) rectangular integrator are uncertain, because the pulse separation was equal to the 100-ms integration time. One could argue that varying the masker intensity within 100 ms of the last pulse would change the output of the integrator and thus affect the threshold. However, one could also argue that the 100-ms pulse separation allowed detection on the basis of independent observations, which could conceivably cause the temporal segment during which the masker varied to be excluded from the decision.

[2]It should be noted that the equivalent rectangular duration of 18.75 ms for a single pulse in the present experiments is slightly more than six times the 3-ms duration of a look assumed by Viemeister and Wakefield (1991). Thus, the detection contributions measured in the present paper may represent averages over several looks.

[3]Buus *et al.*'s (1996) finding that the weights are independent of short-term masker level for single-band and incoherent maskers does not indicate that the detectability of a particular signal pulse is independent of masker level. The weights show how much the detectability changes when the intensity of a pulse changes in a particular situation and are not necessarily dependent on the overall detectability for a particular pulse-number-and-masker condition. Weights that are nearly independent of short-term masker level indicate that some fixed increase in signal energy raises the detectability equally whether the extra energy occurs when the masker level is high or when it is low. In other words, the gain in detectability provided by adding energy to some signal pulse is the same whether the pulse has high or low detectability. The detectability of a particular pulse is not given directly by the conditional psychometric functions, but may be inferred from their intercepts. Recall that the conditional psychometric function shows how the overall d' grows when the level of a particular pulse is varied, and that pulses were selected according to the short-term masker level around them. Thus, the intercept shows the d' that would be obtained if the intensity of pulses occurring within a particular range of short-term masker levels were set to zero. That is, the intercept shows the detectability obtained for five pulses, none of which occurs while the short-term masker level is in the particular range being considered. Buus *et al.*'s (1996) Fig. 1 shows that the intercept increases with increasing short-term level of the masker in the temporal vicinity of the signal pulse under consideration. This finding indicates that performance would be worse if signal pulses were deleted when the masker level is low than if they were deleted when the masker level is high. It follows that signal pulses presented during periods of low masker levels are more detectable than signal pulses presented during periods of high masker levels. In summary, the conditional psychometric functions are characterized by a slope and an intercept. The intercept shows the combined detectability of pulses not being considered. The slope is proportional to the weight on the signal pulse being considered, and shows how much the detectability increases when increasing the intensity of a pulse occurring during a period when the short-term masker level is in a particular range. The finding that the intercept increases with the short-term masker level around a particular pulse shows that pulses occurring during periods of high masker energy contribute the least to detection. The finding that the slopes (weights) are nearly independent of short-term masker level shows that the *gain* in detectability obtained by adding some amount of energy to the signal is the same whether it is added during periods of low masker levels or during periods of high masker levels.

[4]Similar weights are obtained whether it is assumed that d'^2 grows proportionally with I_n^2 (as stated in the derivation from the multiple-looks theory) or that d' grows proportionally to I_n [as followed from Buus *et al.*'s (1996) derivation]. In addition, Buus *et al.* (1996) found that the weights obtained by their method were almost identical to those obtained by Lutfi's (1995) biserial-correlation analysis, which makes no assumptions about the form of the psychometric function. Thus, the weights obtained by the present method depend relatively little on assumptions about the form of the psychometric function.

Berg, B. G. (**1989**). ''Analysis of weights in multiple observation tasks,'' J. Acoust. Soc. Am. **86**, 1743–1746.

Buus, S. (**1997**). ''Auditory masking,'' in *Encyclopedia of Acoustics, Vol. 3*, edited by M. J. Crocker (Wiley, New York), pp. 1427–1445.

Buus, S., Schorer, E., Florentine, M., and Zwicker, E. (**1986**). ''Decision rules in detection of simple and complex tones,'' J. Acoust. Soc. Am. **80**, 1646–1657.

Buus, S., Zhang, L., and Florentine, M. (**1996**). ''Stimulus-driven, time-varying weights for comodulation masking release,'' J. Acoust. Soc. Am. **99**, 2288–2297.

Carlyon, R. P., Buus, S., and Florentine, M. (**1990**). ''Temporal integration of trains of tone pulses by normal and by cochlearly impaired listeners,'' J. Acoust. Soc. Am. **87**, 260–268.

Florentine, M., Fastl, H., and Buus, S. (**1988**). ''Temporal integration in normal hearing, cochlear impairment, and impairment simulated by masking,'' J. Acoust. Soc. Am. **84**, 195–203.

Gilkey, R. H. (**1987**). ''Spectral and temporal comparisons in auditory masking,'' in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 26–36.

Green, D. M., Birdsall, T. G., and Tanner, W. P. J. (**1957**). ''Signal detection as a function of signal intensity and duration,'' J. Acoust. Soc. Am. **29**, 523–531.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (**1984**). ''Detection in noise by spectro-temporal pattern analysis,'' J. Acoust. Soc. Am. **76**, 50–56.

Hughes, J. W. (**1946**). ''The threshold of audition for short periods of stimulation,'' Proc. R. Soc. London, Ser. B **133**, 486–490.

Krumbholz, K., Wiegrebe, L., and Schmid, W. (**1996**). ''Temporal Processing of very brief sounds,'' in *Contributions to Psychological Acoustics.*

Results of the Seventh Oldenburg Symposium on Psychological Acoustics, edited by A. Schick and M. Klatte (Oldenburg University Press, Oldenburg, Germany), pp. 457–464.
Levitt, H. (**1971**). ''Transformed up–down procedures in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.
Lutfi, R. D. (**1995**). ''Correlation coefficients and correlation ratios as estimates of observer weights in multiple-observation tasks,'' J. Acoust. Soc. Am. **97**, 1333–1334.
McFadden, D. (**1986**). ''Comodulation masking release: Effects of varying the level, duration, and time delay of the cue band,'' J. Acoust. Soc. Am. **80**, 1658–1667.
Moore, B. C. J. (**1989**). *An Introduction to the Psychology of Hearing*, 3rd ed. (Academic, London).
Moore, B. C. J., Hall, J. W., Grose, J. H., and Schooneveldt, G. P. (**1990**). ''Some factors affecting the magnitude of comodulation masking release,'' J. Acoust. Soc. Am. **88**, 1694–1702.
Plomp, R., and Bouman, M. A. (**1959**). ''Relation between hearing threshold and duration for tone pulses,'' J. Acoust. Soc. Am. **31**, 749–758.
Schooneveldt, G. P., and Moore, B. C. J. (**1989**). ''Comodulation masking release (CMR) as a function of masker bandwidth, modulator bandwidth, and signal duration,'' J. Acoust. Soc. Am. **85**, 273–281.
Viemeister, N. F. and Wakefield, G. H. (**1991**). ''Temporal integration and multiple looks,'' J. Acoust. Soc. Am. **90**, 858–865.
Wiegrebe, L., Krumbholz, K., Schmid, W., and Schmidt, S. (**1996**). ''Detektion transienter Signale-Ein Maß für Intensitätsintegration im Gehör?,'' Tagungsband DAGA 96, 332–333.

Enhancing the speech envelope of continuous interleaved sampling processors for cochlear implants

Luc Geurts[a] and Jan Wouters[b]
Lab. Exp. ORL, KULeuven, Kapucijnenvoer 33, B 3000 Leuven, Belgium

(Received 17 July 1998; revised 30 October 1998; accepted 8 January 1999)

A new method to code the speech envelope in continuous interleaved sampling (CIS) processors for cochlear implants is proposed. In this enhanced envelope, the rapid adaptation seen in the response of auditory nerves to sound stimuli is incorporated. Two strategies, one using the standard envelope (CIS) and one using the enhanced envelope (EECIS), were tested perceptually with six postlingually deafened users of the LAURA cochlear implant. The tests included identification of stop consonants in three different vowel contexts and monosyllabic consonant–vowel–consonant (CVC) words. Significant improvements in correct identification scores were observed for stop consonants in intervocalic /a/ context ($p=0.026$): average results varied from 46% correct for CIS to 55% for EECIS. This improvement was mainly due to the better transmission of place of articulation. The differences in identification scores for stop consonants in /i/ and /u/ context were not significant. The identification scores for the medial vowels of the CVC words were significantly higher when the EECIS strategy was used: average results increased from 39% correct to 46% correct ($p=0.018$). No significant differences were observed between the results for initial and final consonants of the CVC words. The present results demonstrate that the inclusion of the rapid adaptation in the speech processing for cochlear implants can improve speech intelligibility. © *1999 Acoustical Society of America.* [S0001-4966(99)02704-6]

PACS numbers: 43.66.Ts, 43.72.Ew, 43.71.Ky [JWH]

INTRODUCTION

Cochlear implants electrically stimulate the auditory nerves of deaf patients in order to obtain some level of speech perception. The speech-processing algorithms intend to mimic some of the physiological functions of the cochlea and the auditory-nerve fibers. A widely used strategy is the continuous interleaved sampling (CIS) strategy (Wilson *et al.*, 1991), in which brief pulses are presented to each electrode channel in the cochlea at a high rate of about 1000 pulses per s and in a nonoverlapping sequence. Each pulse train is modulated with the speech envelope in a different frequency band, together covering the speech frequency range. The tonotopic organization (place-frequency coding) of the cochlea is simulated by matching each frequency band to a corresponding electrode channel.

Another physiological mechanism in hearing is the adaptation effect of auditory-nerve fibers. In general, adaptation is the variation in response which occurs during a constant stimulus condition. For auditory nerves, it consists of at least two decaying components: rapid adaptation with a time constant of several milliseconds, and short-term adaptation with a time constant of about 40 ms (Smith, 1979; Westerman and Smith, 1984). As observed at the onsets of speech stimuli, probability of discharge quickly rises to a maximum and then gradually decays to a steady-state level. These ''peaks'' of neural discharge last about 15–20 ms and the more sudden the intensity of the stimulus increases, the more prominent they are (Delgutte, 1980, 1997). Several researchers included this effect in models describing the peripheral auditory processing. Eggermont (1985) developed a model for auditory adaptation, based on existing stochastic models for birth and death processes. In the inner-hair-cell model proposed by Meddis *et al.*, adaptation is caused by the depletion of neurotransmitter at the synapses between hair cells and auditory nerves (Meddis, 1986, 1988; Meddis *et al.*, 1990; Lopez-Poveda *et al.*, 1997). Patterson's time-domain model of peripheral auditory processing includes adaptive thresholding (Patterson *et al.*, 1995; Holdsworth and Patterson, 1993). Delgutte measured and modeled the modulation transfer functions of several auditory nerves differing in characteristic frequency and used this model to describe the neural encoding of the temporal envelope of speech (Delgutte, 1997). The temporal variations of the speech envelope at a syllabic rate seem to be emphasized by these adaptation peaks. These low-frequency modulations are probably very important for speech intelligibility. Drullman *et al.* (1994) showed that reducing the slow temporal modulations of the envelope leads to poor speech recognition. Furthermore, speech with reduced spectral information, but with the envelope information preserved, is still intelligible (Shannon *et al.*, 1995).

Since one of the origins of the adaptation effect is at the synapses between inner hair cells and auditory nerves, and since a cochlear implant bypasses the hair cells, a different effect is seen in the nerve fiber's response to electrical stimulation. It is still true that the discharge rate may decay during the stimulus, which indicates that adaptation is not solely determined by the properties of the hair-cell/auditory-nerve synapse (Wilson *et al.*, 1995; Killian *et al.*, 1994; Dynes and Delgutte, 1992; Parkins, 1989; Javel *et al.*, 1987; van den

[a]Electronic mail: Luc.Geurts@uz.kuleuven.ac.be
[b]Electronic mail: Jan.Wouters@uz.kuleuven.ac.be

Honert and Stypulkowski, 1987). But, the effect differs in several respects. The adaptation in response to electric stimulation has a larger time constant and a much smaller amplitude, compared to acoustic stimulation. Moreover, the effect is not consistent, since not all fibers seem to show a decay in discharge rate (Dynes and Delgutte, 1992; Killian *et al.*, 1994). On the other hand, some fibers can cease to respond after a short time period of activity. This is the result of depolarization block and occurs mostly for stimuli at high rates and intensities (Killian *et al.*, 1994; Javel *et al.*, 1987). In conclusion, the adaptation effect should be added to a cochlear prosthesis processor to better mimic the neural response patterns to normal acoustic stimulation.

In this report, we present a method to incorporate this effect in the speech processing for cochlear implants, using a new algorithm to calculate the amplitude of the speech envelope. The high rate of stimulation of CIS processors allows the enhancement of peaks in the envelope with a duration of 10–20 ms. Two strategies, one using the standard envelope (CIS) and one using the enhanced envelope (EECIS), were tested perceptually with users of the LAURA cochlear implant[1] (Peeters *et al.*, 1993).

I. SPEECH PROCESSING

The two speech-processing algorithms in this study consist of the following steps: prefiltering, bandpass filtering, envelope extraction, instantaneous compression, current mapping, and pulse-train modulation. The algorithms differ in the way the speech envelope is extracted; the other steps are identical. The processing for one frequency band is shown in Fig. 1: the left, bold trace in the flow diagram corresponds to the CIS algorithm, while the whole diagram corresponds to the EECIS algorithm.

A. Prefiltering and bandpass filtering

The input of the speech-processing algorithm is digitized speech at a sample rate of 10 kHz and a resolution of 16 bit. Prefiltering is applied in order to compensate the typical slope of −6 dB in the long-term average speech spectrum, starting from about 500 Hz (Byrne *et al.*, 1994). The preemphasis filter is a first-order derivative filter, with the 3-dB cutoff frequency at 700 Hz, and a gain at dc of 0 dB.

Next, the signal is filtered by different bandpass filters covering the speech frequency range. The number corresponds to the number of active electrode channels in the cochlea of the implanted subject. The cutoff frequencies are linearly spread below 1 kHz and logarithmically above 1 kHz. For 8 channels, these are 100, 325, 550, 775, 1125, 1634, 2372, 3444, and 5000 Hz, and for 7 channels, these are 100, 357, 614, 934, 1421, 2161, 3287, and 5000 Hz.

B. CIS: Standard envelope extraction

The remaining processing is identical for each channel and is shown for one channel in Fig. 1. The standard envelope is extracted in the left, bold trace. First, the signal is half-wave rectified, in contrast with the full-wave rectification originally used in the CIS strategy (Wilson *et al.*, 1991). This is done because in the latter case the frequency, and thus the rate pitch, is doubled in the low-frequency channels. This is not the case for the high-frequency channels, since in that case the pitch appears as an envelope modulation and not as a single harmonic.

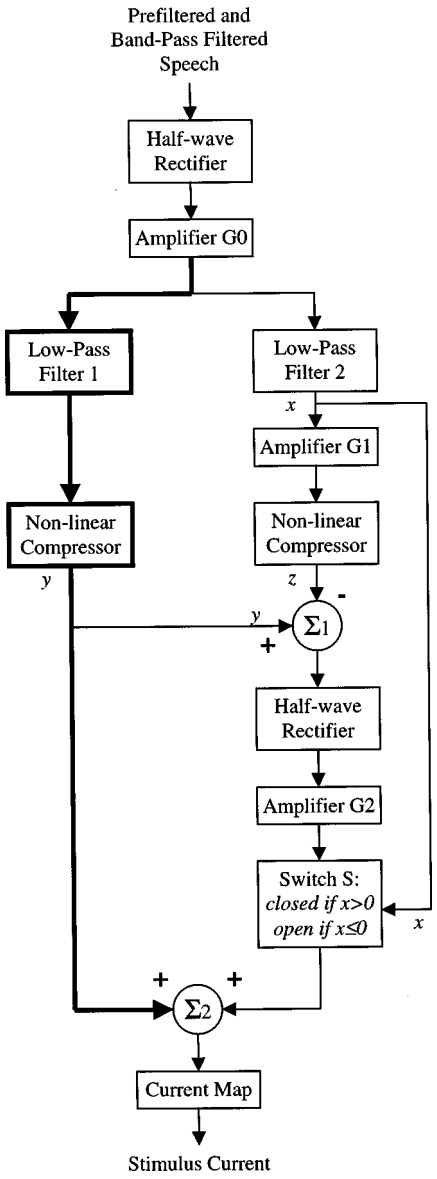


FIG. 1. Processing scheme for the standard envelope (left, bold trace) and the enhanced-envelope extraction (both traces).

Second, the speech envelope is smoothed by low-pass filtering the rectified signal (left, bold trace, low-pass filter 1: Butterworth 4th order, cutoff frequency at 400 Hz). The resulting signal is amplified with a channel-specific gain G0 to compensate for the greater attenuation caused by rectification in high-frequency channels.

Third, the smoothed envelope is instantaneously compressed using the nonlinear function shown in Fig. 2. This compression function is identical to the one used in the subjects' clinical speech processors and consists of a piecewise linear approximation of a square root curve. Sample values above 2^{14} are clipped. However, if the speech signal at the input has a resolution of 16 bit (values between -2^{15} and 2^{15}), only a very small number of samples exceed the value 2^{14} after bandpass filtering and smoothing.

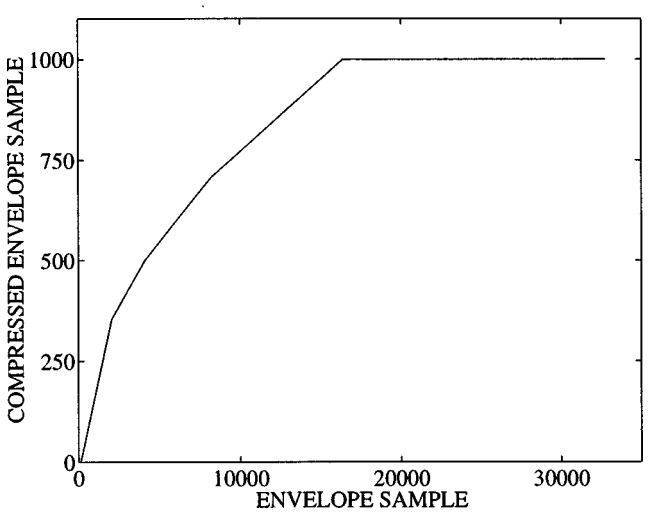


FIG. 2. Piecewise linear compression function. The input of the function (horizontal axis) is the 16-bit sample value of the envelope. The output is an integer value between 0 and 1000.

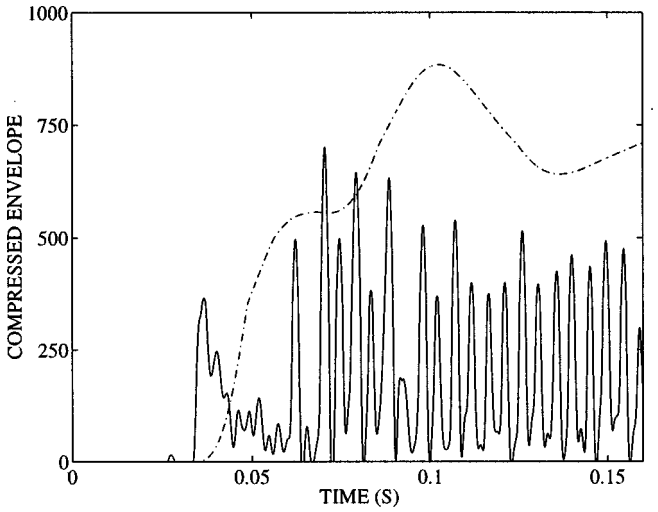


FIG. 3. Compressed envelopes of syllable /ka/ in channel 8 (frequency band between 3444 and 5000 Hz) in function of time. The solid line is the envelope of the fast filter (cutoff frequency at 400 Hz) and corresponds to signal *y* in Fig. 1. The dashed line is the amplified output of the slow filter (cutoff frequency at 20 Hz) and corresponds to signal *z* in Fig. 1.

C. EECIS: Enhanced envelope extraction

The enhanced envelope is the sum of the standard envelope extracted in the left, bold trace of Fig. 1 and a peak signal, extracted in the right trace whenever there is a sudden increase in the intensity of the input signal. This implementation has the advantage that the only difference between both strategies is the presence of the peak signal. So, differences in performance are completely due to this peak signal. Another advantage is that only one extra parameter is needed to ''fit'' the algorithm to the cochlear implant user, namely, the height of the peak signal. Rapid adaptation is thus simulated by increasing the amplitude of the speech envelope at sudden intensity increases.

The first step in the right trace of Fig. 1 is the derivation of another envelope from the rectified signal, using a low-pass filter with a much lower cutoff frequency (signal *x*, right trace, low-pass filter 2: Butterworth 4th order, cutoff frequency at 20 Hz). This envelope is amplified and compressed, using the same compression function as the one in the left, bold trace. An example of the two envelopes is shown in Fig. 3. The solid curve shows the compressed envelope *y* extracted in the left, bold trace. The dashed curve shows the envelope *z* extracted in the right trace. The speech signal is the syllable /ka/ (male speaker) in channel 8, so the bandpass filtering has cutoff frequencies of 3444 and 5000 Hz. The burst of the /k/ occurs between 0.03 and 0.05 s; the /a/ starts at 0.06 s.

The lower cutoff frequency of low-pass filter 2 serves two goals. First, more high-frequency components are rejected, leading to a smaller ripple. It is clear from the example that no F0 modulation is present anymore in the dashed envelope. Second, the filter has a bigger delay, so the output of this filter lags the output of the other one. The amplification G1 is chosen such that the ''slow'' envelope *z* is lying above the ''fast'' envelope *y* for the stationary part of the signal (the vowel). A suitable value is an amplification factor of 6, for all channels. Using this value, the criterion was met for all speech tokens in a large set consisting of consonants in three different vowel contexts /a/, /i/, and /u/ and of vowels in /pVt/ context, recorded from two male and two female speakers.

However, at the onset of the speech utterance, the fast envelope *y* will always be above the slow one (*z*). This part is the peak signal to be extracted. This is done by subtracting the slow envelope from the fast one ($\Sigma 1$ in Fig. 1), thereby only retaining the positive values using half-wave rectification. The duration of the peak is on the order of magnitude of 10 ms, which is in correspondence with the time constant values of the rapid adaptation as measured by Delgutte (1980). The extracted peak signal is amplified by a factor G2 between 1 and 2, the exact value depending on a subjective evaluation by each subject (see further).

The enhanced envelope is the summation of the standard speech envelope, extracted in the left trace, and the peak signal, extracted in the right trace ($\Sigma 2$ in Fig. 1). An example of the standard envelope (middle plot) and of the enhanced envelope (lower plot) in all channels for the syllable /ka/ is shown in Fig. 4. Notice the difference in amplitude for the envelope of the burst. In general, the enhanced envelope emphasizes the onset of the typical intensity increases in speech.

The switch *S* in Fig. 1 is used to prevent ''false peaks.'' These can occur when there is a rapid decay in intensity, immediately after a sudden increase in intensity. In that case, the ringing caused by the low-pass filtering leads to a temporal negative output, even while the input remains positive. This effect is more prominent for the output *x* of the slow low-pass filter 2, since it has a larger time constant. This results in false peaks: the fast envelope *y* of the left, bold trace will be above the slow envelope *z* of the right trace, only because the latter is negative. By opening switch *S* when the output *x* of low-pass filter 2 is negative, these false peaks will not be added.

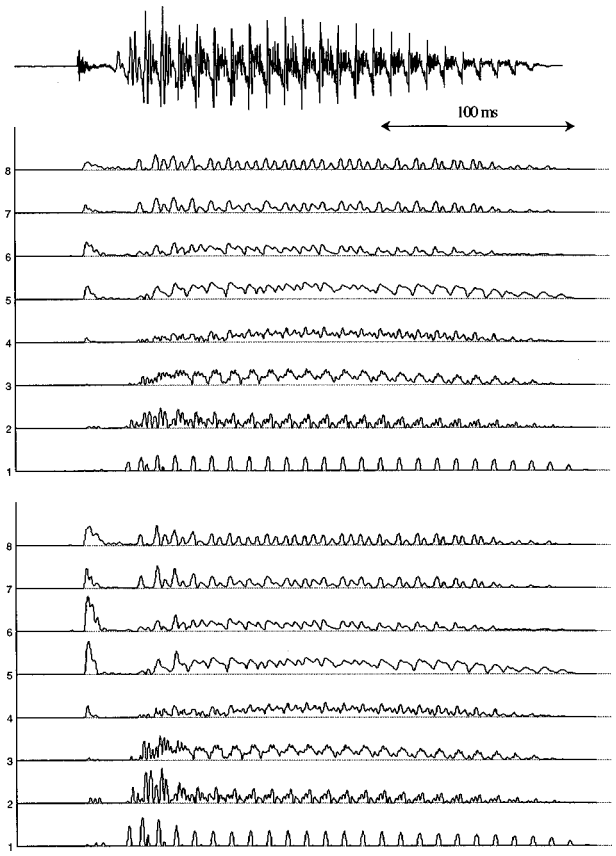


FIG. 4. The upper plot shows the sampled wave form of the syllable /ka/ as a function of time. The middle plot shows the corresponding envelopes in 8 channels for the CIS algorithm. The lower plot shows the envelopes for the EECIS algorithm. The value of G2 equals 1.5. The channel numbers are indicated before each envelope trace. Channel 1 corresponds to the most apical channel, channel 8 to the most basal channel.

D. Current mapping and pulse modulation

The envelope samples are mapped to current values (last block at the bottom of Fig. 1), taking the dynamic range of the channel of the subject under test into account. A linear function is used, which maps an envelope sample value of 1 to the current threshold, and a value of 1000 to the most comfortable current level. The following parameters correspond to the standard strategy used in the LAURA cochlear implant. The current envelopes in all channels are modulated with pulse trains with an overall rate of 10 000 pps (pulses per second). If 8 channels are used, the pulse rate per channel is 1250 pps, while this rate equals 1429 pps in the 7-channel case. The width of each pulse is constant and equal to 40 μs/phase.[2] Following the CIS strategy, the pulse trains of all channels are interleaved. The channel order of stimulation is staggered: 1, 4, 7, 2, 5, 8, 3, and 6, where channel 1 corresponds to the most apical channel and channel 8 to the most basal channel. Stimulation on the nonactive electrode channel is skipped for subjects using only 7 channels.

E. Implementation

All the speech tokens used for the perceptual evaluation are processed off-line and presented electrically to the subject. The envelope value and the peak value before current mapping, and the channel number of each electrical pulse are stored in files on hard disk or CD-ROM. During the test, these files are read by customized software, and the values are sent to a digital signal processor (DSP) board connected to the computer. The DSP program requires subject-dependent values for the following calculations: the amplification G2 of the peak signal, which will be added on the envelope signal, and the mapping of the envelope values to current values within the dynamic range. The advantage of working this way is that the same files, containing the processed words, can be used for all subjects. The DSP program also serves as a sequencer, generating the time code to drive the internal part of the cochlear implant device. The DSP board is connected to the external coil of the implant system of the subject. Communication with the internal part is accomplished by inductive coupling with the internal coil.

II. IDENTIFICATION EXPERIMENTS

Different speech tests are performed in order to compare the two algorithms, the standard envelope extraction (CIS) and the enhanced envelope extraction (EECIS).

TABLE I. Some details of the eight subjects who took part in this experiment. Columns 7 and 8 show the results for an identification task of ten vowels in /h/-vowel-/t/ context, and 16 consonants in /a/-consonant-/a/ context, respectively (van Wieringen and Wouters, 1999). A suitable value for G2 (last column) could not be found for subjects HS and JH (see the text).

Subject	Age (yrs)	Duration of profound deafness (yrs)	Etiology	Implant experience (yrs)	Number of active electrode channels	Vowel score (%)	Consonant score (%)	Value of *G*2
AP	60	3	cochl. otosclerose	1	8	49	29	1.5
DD	32	1	trauma	3	8	61	35	1
GG	69	14	meningitis	10	7	17	28	1.5
SV	29	6	unknown	4	8	73	58	1.5
WM	48	5	progressive	1	8	49	51	1.5
AE	42	3	Menière	2	7	79	51	2
HS	44	12	congenital	2	8	40	35	⋯
JH	16	3	meningitis	3	8	53	45	⋯

A. Subjects

Eight postlingually deafened users of the LAURA cochlear implant participated in this experiment. Table I lists some relevant information about each subject. All subjects use the LAURAFLEX speech processor, in which the CIS strategy is implemented. Columns 7 and 8 show the results for vowel and consonant identification tasks obtained in a previous study (van Wieringen and Wouters, 1999).

B. Fitting

Thresholds (THR) and most-comfortable levels (MCL) were obtained at the beginning of the first session and were identical for both strategies. While for the CIS strategy the MCL value corresponds to the maximum possible pulse amplitude, this is not the case for the EECIS strategy: due to the addition of the peak signal, the envelope can temporarily exceed the MCL value. The processing schemes are designed such that the full dynamic range is used if the input speech samples are within the full 16-bit range. Therefore, the loudness of the words was always at a comfortable level for all subjects, and no correction of the overall stimulation level was needed.

The height of the peak signal is determined by the value of amplification factor G2. This gain is estimated for each subject by a a subjective procedure as follows. Three voiceless stop consonants (p,t,k) in intervocalic /a/ context of a male speaker, processed by both strategies, were presented to the subject. Voiceless stop consonants in /a/ context were used since, for these phonemes, a large peak signal was prominent in all channels. If the value of G2 was too low, a difference could not be heard between both strategies for any of the words. If the value was too high, the word, processed by the EECIS strategy, sounded distorted: the peak signal sounded as a short, loud, nonspeechlike crack. For six subjects, a suitable value could be found, varying between 1 and 2 (see Table I). They reported that the words with the enhanced envelope sounded more clear, distinct, and intelligible, as if they were better articulated. The other subjects, HS and JH, represent two extreme cases. Subject HS could not hear any difference, even when applying the technically highest-possible currents in the peak signals. For subject JH, the words with the enhanced envelope always sounded distorted. The height of the peak signal had to be made so small that he could not hear a difference anymore. Both subjects were excluded from further testing. Possible reasons for this effect are not obvious.

C. Speech material and tests

Two speech sets were used: stop consonants in vowel context and monosyllabic words of the form consonant–vowel–consonant (CVC). All the words within each set were balanced in rms level. The first set included nonsense words of the form vowel–consonant–vowel (VCV) where V={/a/, /i/,/u/} and C={/p/,/t/,/k/,/b/,/d/,/g/}. Stop consonants were used, since the effect of the algorithm incorporating the rapid adaptation is largest for these phonemes compared to others. This is also the case in normal hearing, where large adaptation peaks are seen at, or shortly after, stop consonants (Delgutte, 1997). All the words were pronounced by two male and two female speakers. Different contexts and different speakers were used to test the robustness of the new algorithm. Presentation was in blocks of 51 words: six stop consonants of four speakers were presented twice in each block, plus three randomly chosen words at the beginning of the block, which were not included in the analysis. Within each block, the vowel context and the processing strategy were the same. Twelve blocks were run for each subject in an ABAB design: each strategy was used in every other block. At the start of the session, the subjects could listen to the six stop consonants in intervocalic /a/ context of one male speaker, processed by both strategies. This way, they could get familiar with the words in each strategy. Since after testing no significant differences in identification scores were observed between these ''training'' words and the others used in the test (other speakers and other contexts), all responses were included in the analysis. Response alternatives for the identification task were displayed on a graphics tablet and the listeners were required to indicate the alternative with the tablet pen to register their response. Feedback was not given. Analysis consisted of calculation of percentage-correct responses and construction of confusion matrices. The latter permits the calculation of relative feature information transmission (Miller and Nicely, 1955). Two features were used: VOICING, with categories *unvoiced* {p,t,k} and *voiced* {b,d,g}, and PLACE OF ARTICULATION, with categories *front* {p,b}, *middle* {t,d} and *back* {k,g}.

In order to evaluate both strategies in a wider context, the subjects also completed an identification task with monosyllabic CVC words. These were part of the NVA list (Wouters *et al.*, 1994), which consists of equivalent sublists of 12 Dutch/Flemish words. These words contain little redundant information: perceptual experiments revealed that the number of statistically independent phonemes varied between 2.4 and 2.7 for normal hearing and hearing-impaired subjects (Bosman *et al.*, 1992). The words were recorded from one male speaker and six sublists were presented for each strategy. The response for the first word in each sublist was not included in the analysis. Percent-correct phoneme scores were calculated for initial consonant, medial vowel, and final consonant separately. For each of these groups, confusion matrices were constructed. The matrices of the six subjects were added together and feature analysis was performed (Miller and Nicely, 1955). For the consonants, the features were MANNER (six categories), PLACE (three categories) and VOICING (two categories). For the vowels, the features were DURATION, first formant frequency (F1), and second formant frequency (F2), all having three categories.

D. Results

1. Stop consonants

Table II gives an overview of the results of the stop consonant identification task, for each vowel context separately. The results averaged across all subjects are also shown in Fig. 5. The biggest effect of the new algorithm was observed for stop consonants in intervocalic /a/ context, for which there was an absolute increase in identification score

TABLE II. Percentage-correct responses of each subject for the stop consonant identification task. The results are given for each vowel context separately. The last row shows the results, averaged across all subjects.

	aCa		iCi		uCu	
	CIS	EECIS	CIS	EECIS	CIS	EECIS
AP	26	46	26	43	32	40
DD	36	51	34	30	43	39
GG	39	46	39	35	47	45
SV	46	46	64	67	40	32
WM	57	63	49	58	36	40
AE	73	80	67	73	56	47
Average	46	55	47	51	42	41

TABLE III. Percentage-correct responses of each subject for the CVC identification task. The results are shown for the initial consonant, medial vowel, and final consonant separately. The last row shows the results, averaged across all subjects.

	Initial consonant		Medial vowel		Final consonant	
	CIS	EECIS	CIS	EECIS	CIS	EECIS
AP	50	50	48	61	44	47
DD	39	52	24	27	35	29
GG	36	42	15	17	33	33
SV	47	44	21	35	38	35
WM	61	58	50	58	56	61
AE	61	65	73	77	68	62
Average	49	52	39	46	46	44

of 9%. The effect was smaller for the /i/ context (+4%) and there was a small negative effect for the /u/ context (−1%). For each context, the results for the two algorithms were compared in a paired samples *t*-test:[3] the difference in identification scores was significant for the /a/ context ($t[5]=3.15$, $p=0.026$), but not for the /i/ context ($t[5]=1.37$, $p=0.23$) and the /u/ context ($t[5]=-0.67$, $p=0.53$). The analysis of relative feature information transmission revealed that the two features PLACE OF ARTICULATION and VOICING contribute to the better performance for the EECIS algorithm in the /a/ context. The relative transmission of the feature PLACE OF ARTICULATION, averaged across the data of all subjects, varied from 21% for the CIS strategy to 29% for the EECIS strategy. The relative transmission of the feature VOICING increased from 58% to 61%. Since this measure is more sensitive for small numbers, it can be concluded that the better performance is mainly due to the better transmission of the feature PLACE OF ARTICULATION. This is very encouraging, since previous studies have shown that this feature is very difficult to perceive for cochlear implantees (Tyler and Moore, 1992; Dorman *et al.*, 1990).

2. CVC words

A summary table of the identification scores for the CVC words test is presented as Table III. The scores, averaged across all subjects, and the percentages-transmitted information of all features are shown in Fig. 6. Only small differences between the results of both processing schemes are observed for the consonants. The 3% increase in percentage-correct responses for the initial consonants is not significant ($t[5]=1.12$, $p=0.31$); neither is the 2% decrease for the final consonants ($t[5]=-0.62$, $p=0.56$). The new strategy does not seem to have a notable effect on the information transmission of any feature. On the other hand, identification of the medial vowel is significantly improved ($t[5]=3.45$, $p=0.018$): average scores increased from 39%

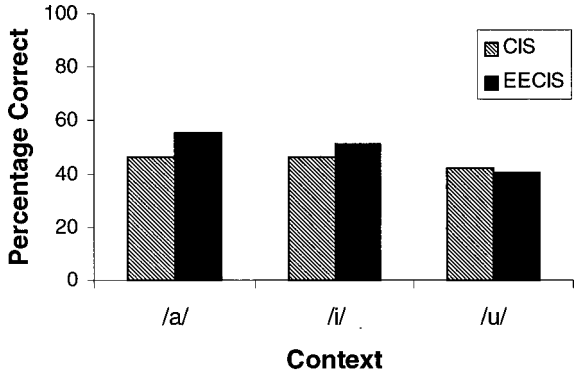


FIG. 5. Percentage-correct stop consonant identification for the vowel contexts /a/, /i/, and /u/, and for the algorithms CIS and EECIS.

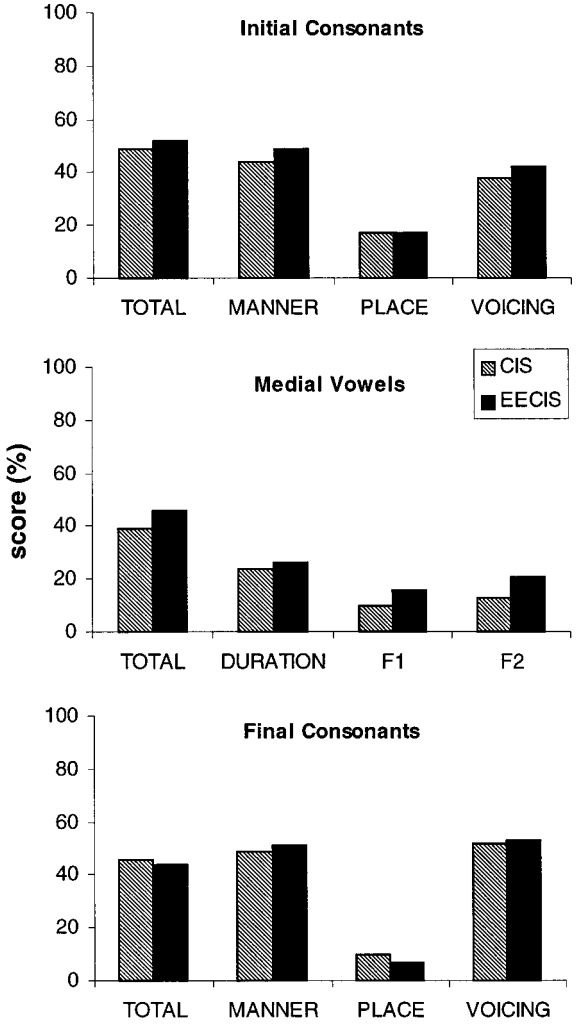


FIG. 6. Percentage-correct phoneme identification (TOTAL) and relative transmitted information of the features for the initial consonants (upper plot), medial vowels (middle plot), and final consonants (bottom plot) of CVC words, for the algorithms CIS and EECIS.

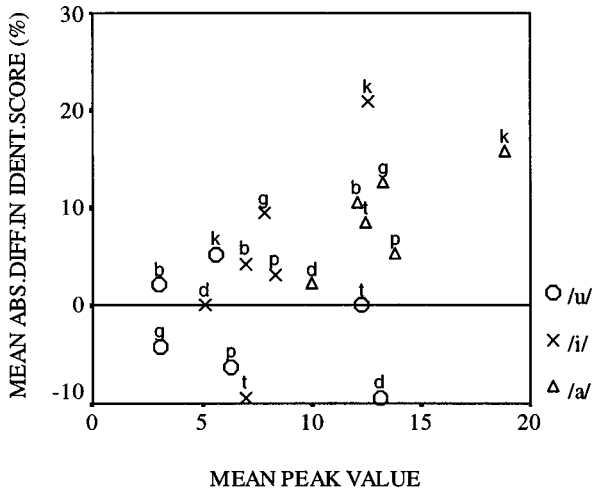


FIG. 7. Absolute difference in identification score between both algorithms, averaged across all listeners, in function of the mean peak value for the six consonants in the three vowel contexts, averaged across all speakers.

correct for the CIS strategy to 46% correct for the EECIS strategy. This was mainly due to the better transmission of the formant frequencies: from 10% to 16% for the first formant frequency, and from 13% to 21% for the second formant frequency. This is a surprising result, since only the onsets of the vowels are enhanced in each of the 7 or 8 channels, and only if the preceding consonant has less energy than the vowel in that channel.

III. SIGNAL ANALYSIS

The peak signals were further investigated to find possible underlying reasons for the observed differences in improvement for the different vowel contexts. The peak signals were extracted separately by the EECIS algorithm, so the mean peak values at the onset of the second syllable (stop vowel) could easily be calculated. The mean peak value is defined as the average value of all nonzero peak samples across all channels. Mean peak values were calculated for the 8-channel and the 7-channel algorithm. The correlation between the obtained measures for both cases was very high ($r=0.997$, $p<0.001$), so only the data for the 8-channel algorithm were retained for further analysis. One-way analysis of variance on the factor SPEAKER also revealed that there were no significant differences in the mean peak value of each syllable across the four speakers ($F[3,51]=0.23$, $p=0.88$). Moreover, all correlations between the mean peak values for all possible pairs of speakers were significant at the 0.01 level, and had values between 0.6 and 0.7. This indicates that the peak signal is not dependent on the speaker, which means that the EECIS algorithm is robust with respect to differences in pronunciation. For further analysis, the mean peak values were averaged across all speakers.

Figure 7 illustrates the relationship between the mean peak value (horizontal axis) averaged across all speakers, and the mean absolute difference in identification scores, averaged across all listeners (vertical axis) for each stop consonant in each vowel context. The ordering on the horizontal axis reveals that the mean peak values are highest for the /a/ context and lowest for the /u/ context, and those for the /i/ context lie in between. One explanation is that peak values are averaged across all channels. In the case of the /a/ context, peaks are prominent in all channels, leading to a large average value. On the other hand, only a limited number of channels contain peak signals when the vowel context was /i/ (lowest and highest channels) or /u/ (only lowest channels). The orderings of the mean peak values for the different stops are consistent with the data of Repp and Lin (1989) for English stops. They also found the highest onset amplitudes for velar stop consonants preceding /a/ or /i/, while alveolar stop consonants had the highest onset amplitudes if they were followed by /u/.

In Fig. 7, 14 out of the 18 tokens have non-negative values on the vertical axis, which means that for these tokens identification scores for the EECIS algorithm are at least the same, but mostly higher, compared to those of the CIS algorithm. There is a moderate, but significant, positive correlation between the mean peak value for each token and the difference in identification score between both algorithms for the corresponding token ($r=0.51$, $p<0.05$). Within each context separately, correlations were 0.78, 0.80, and -0.41 for the /a/, the /i/, and the /u/ context, respectively. The negative correlation for the /u/ context possibly reflects the confusion that arises, since alveolar stop consonants had the largest peak signals, as opposed to the /a/ and /i/ context.

IV. DISCUSSION

The proposed technique to enhance the speech envelope is based on the rapid adaptation seen in the neural discharge rate at the onset of a sound stimulus. However, it is very difficult to ensure that quantitatively the same effect in the neural response is obtained. One way would be to make single fiber recordings of the response of electrically stimulated auditory nerves in animals to pulse-train stimuli with and without peak signals. Another way would be to measure evoked potentials in humans with the same stimuli. On the other hand, it can be assumed that, at least qualitatively, the desired effect is obtained, since increasing the current amplitude increases the discharge probability of the electrically stimulated nerve fibers. A side effect of this current increase is that there is also a corresponding recruitment of more fibers with an increase in the synchronization between them. This is probably unlike the events occurring in normally adapting fibers. An alternative technique would be to increase the stimulation rate instead of the amplitude of the pulses. However, the current technology of available cochlear implant systems does not yet allow the necessary flexible variations in pulse rate over a large interval.

It is difficult to determine the value of amplification factor G2 (height of the peak), in order to create a similar effect as in normal hearing. Since an objective, quantitative measure was not available, a subjective measure was used, i.e., the response of the subject. The advantage of the proposed technique is that the peak signal is extracted separately from the standard speech envelope, so a peak signal, ''fitted'' to the subject, can be added to the standard speech envelope. For each subject, the interval of suitable values for G2 was rather small: the ratio between the smallest value, resulting

in an effect, and the largest value, resulting in a distorted sound, was not larger than 2. Across all subjects, the ratio between the smallest and the largest value used in the experiments was also 2. The value of G2 was kept the same across all channels, which makes the proposed technique suitable for clinical use. A possible improvement of the strategy is to make this value channel dependent.

The fitting of the peak signal before testing also revealed that not all subjects benefit from the new algorithm. One subject could not hear any difference when a peak signal was added, while for another subject the word always sounded distorted in that case. We could not find an explanation for this. It is possible that the electrical–neural interface of these subjects differs from those of the other subjects, caused by, e.g., differences in the number and distribution of surviving neurons, physiologic function of these neurons, and distance between stimulating electrodes and these neurons.

An objection might be that the improvements in identification of stop consonants are due to the amplification of the bursts. In other words, the same results could possibly be obtained if the words were simply amplified. There are, however, three reasons to refute this. First, the enhanced envelope extraction technique only amplifies the onsets and thus performs a selective amplification. In the case of the stop consonants, the bursts are amplified, relative to the following vowel. It is shown that increasing the consonant–vowel intensity ratio improves consonant recognition by hearing-impaired listeners (Kennedy *et al.*, 1998). The proposed envelope-extraction technique can be seen as a method to obtain this effect. Second, adaptation can also be seen as selective amplification. Plots of neural activity show that the discharge rate for stop consonants is higher than what might be expected from their energy (Delgutte, 1997). Third, it is not feasible to increase the stimulation level for the CIS scheme to match that of the consonants in the EECIS scheme. A peak amplification factor G2 of 2 means that the onset is amplified with a factor 3. If all current pulses were amplified by this factor, the loudness of the speech stimuli would be unbearable.

The obtained results for stop consonants are context dependent. The average improvement in percentage-correct identification was largest for stops in /a/ context, smaller for stops in /i/ context, and slightly negative for stops in /u/ context. This effect is probably due to the fact that the average values of the peaks are largest for stops in /a/ context and smallest for stops in /u/ context, in part caused by the different number of channels containing prominent peak signals. This was also the reason why stops in /a/ context were used in the fitting procedure, since the purpose of the fitting was to determine the maximum possible value of the peak amplification factor G2. A different fitting method for the other contexts could compensate for the context-sensitive outcomes, but it is not feasible to incorporate this in a real-time speech processor, since the following vowel is not known *a priori*. The problem can partly be overcome by applying a compression on the peak signals, making the observed differences in mean peak value smaller. However, compression will obviously have no effect in channels containing no peak signals (/i/ and /u/ context). Another approach is to increase the value of the peak amplification G2, even if this makes the speech signal sound unnatural. Care has to be taken though, since performance can drop if the consonant–vowel intensity ratio is increased too much (Kennedy *et al.*, 1998).

The results for the CVC words are more difficult to interpret. No significant differences were seen between the results for initial consonants, nor for final consonants. There are several possible explanations for this. First, the consonant set consisted of more than only stop consonants: 38% of the initial consonants and 30% of the final consonants were stops. Second, the medial vowel was obviously not always /a/, while the best results were obtained for identification of stop consonants in /a/ context. Third, the subjects did not focus their attention entirely on the consonants alone when they identified the words. On the other hand, their performance in vowel identification increased, mainly due to the better transmission of the formant frequencies. Presumably, the peak signals enhanced the spectral contrast between the preceding consonant and the vowel, analogous to the adaptation effect (Delgutte, 1997).

In general, there was a significant improvement, but it was rather small. Yet, the obtained results were immediate, requiring no training and no feedback. Greater improvements might be achieved if the subjects could use the new processing scheme for an extended time period. Real-time implementation in a wearable speech processor is feasible, given the low complexity of the algorithm.

In summary, a new method to code the speech envelope in CIS processors for cochlear implants is proposed, which incorporates the rapid adaptation seen in the response of auditory nerves to sound stimuli. Significant improvements in speech intelligibility were observed in some conditions, while there was never a significant decrease in the transmission of any of the speech features. This suggests that some perceptual advantage is gained by including rapid adaptation in the speech processing for cochlear implants.

ACKNOWLEDGMENTS

The authors wish to thank all the subjects, who were implanted at the University ENT-Department of the A.Z. St. Augustinus Hospital in Antwerp, for taking part in the experiments. We also thank A. van Wieringen and two anonymous reviewers for their comments on an earlier version of the manuscript. We would also like to thank S. Peeters for his continuing interest in this study. This study was supported by Philips Hearing Implants, the Flemish Institute for the Promotion of Scientific–Technological Research in Industry, and the Fund for Scientific Research–Flanders (Belgium).

[1]LAURA is a registered trademark of Philips Hearing Implants.
[2]For some other users of the LAURA device, pulse widths can be 100 μs/phase or even 200 μs/phase.
[3]Statistical analyses were performed using SPSS 7.5 for Windows.

Bosman, A., Smoorenburg, G. F., and Bronkhorst, A. W. (**1992**). ''Relations between phoneme scores and syllable scores for normal-hearing and hearing-impaired subjects,'' in *The Auditory Processing of Speech, From Sounds to Words*, edited by M. E. H. Schouten (Mouton de Gruyter, Berlin, New York, 1992), pp. 357–362.

Byrne, D., Dillon, H., Tran, K., *et al.* (**1994**). ''An international comparison of long-term average speech spectra,'' J. Acoust. Soc. Am. **96**, 2108–2120.

Delgutte, B. (**1980**). ''Representation of speech-like sounds in the discharge patterns of auditory-nerve fibers,'' J. Acoust. Soc. Am. **68**, 843–857.

Delgutte, B. (**1997**). ''Auditory neural processing of speech,'' in *The Handbook of Phonetic Sciences*, edited by W. J. Hardcastle and J. Laver (Oxford, Blackwell, 1997), pp. 507–538.

Delgutte, B., Hammond, B. M., and Cariani, P. M. (**1997**). ''Neural encoding of the temporal envelope of speech: Relation to modulation transfer functions,'' in *Psychophysical and Physiological Advances in Hearing: Proceedings of the 11th International Symposium on Hearing, Grantham, U.K.*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Wurr, London, 1997), pp. 595–603.

Dorman, M., Soli, S., Dankowski, K., Smith, L., 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.

Drullman, R., Festen, J. M., and Plomp, R. (**1994**). ''Effect of temporal envelope smearing on speech perception,'' J. Acoust. Soc. Am. **95**, 1053–1064.

Dynes, S. B. C., and Delgutte, B. (**1992**). ''Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea,'' Hearing Res. **58**, 79–90.

Eggermont, J. J. (**1985**). ''Peripheral auditory adaptation and fatigue: A model oriented review,'' Hearing Res. **18**, 57–71.

Holdsworth, J. W., and Patterson, R. D. (**1993**). ''Analysis of waveforms,'' UK Patent No. GB 2-234-078-B, UK Patent Office, London.

Javel, E., Tong, Y. C., Shepherd, R. K., and Clark, G. M. (**1987**). ''Responses of cat auditory-nerve fibers to biphasic electrical current pulses,'' Ann. Otol. Rhinol. Laryngol. **96**, Suppl. 128, 26–30.

Kennedy, E., Levitt, H., Neuman, A. C., and Weiss, M. (**1998**). ''Consonant–vowel intensity ratios for maximizing consonant recognition by hearing-impaired listeners,'' J. Acoust. Soc. Am. **103**, 1098–1114.

Killian, M. J. P., Klis, S. F. L., and Smoorenburg, G. F. (**1994**). ''Adaptation in the compound action potential response of the guinea pig VIIIth nerve to electric stimulation,'' Hearing Res. **81**, 66–82.

Lopez-Poveda, E. A., O'Mard, L. P., and Meddis, R. (**1997**). ''A revised computational inner hair cell model,'' in *Psychophysical and Physiological Advances in Hearing: Proceedings of the 11th International Symposium on Hearing, Grantham, U.K.*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Wurr, London, 1997), pp. 112–121.

Meddis, R. (**1986**). ''Simulation of mechanical to neural transduction in the auditory receptor,'' J. Acoust. Soc. Am. **79**, 702–711.

Meddis, R. (**1988**). ''Simulation of auditory–neural transduction: Further Studies,'' J. Acoust. Soc. Am. **83**, 1056–1063.

Meddis, R., Hewitt, M. J., and Shackleton, T. M. (**1990**). ''Implementation details of a computational model of the inner hair-cell/auditory-nerve synapse,'' J. Acoust. Soc. Am. **87**, 1813–1816.

Miller, G. A., and Nicely, P. E. (**1955**). ''An analysis of perceptual confusions among some English consonants,'' J. Acoust. Soc. Am. **27**, 338–352.

Parkins, C. W. (**1989**). ''Temporal response patterns of auditory nerve fibers to electrical stimulation in deafened squirrel monkeys,'' Hearing Res. **41**, 137–168.

Patterson, R. D., Allerhand, M. H., and Giguère, C. (**1995**). ''Time-domain modelling of peripheral auditory processing: a modular architecture and a software platform,'' J. Acoust. Soc. Am. **98**, 1890–1894.

Peeters, S., Offeciers, F. E., Kinsbergen, J., Van Durme, M., Van Enis, P., Dykmans, Ph., and Bouchataoui, I. (**1993**). ''A digital speech processor and various encoding strategies for cochlear implants,'' Prog. Brain Res. **97**, 283–291.

Repp, B. H., and Lin, H. (**1989**). ''Acoustic properties and perception of stop consonant release transients,'' J. Acoust. Soc. Am. **85**, 379–396.

Shannon, R. V., Zeng, F., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). ''Speech recognition with primarily temporal cues,'' Science **270**, 303–304.

Smith, R. L. (**1979**). ''Adaptation, saturation, and physiological masking in single auditory-nerve fibers,'' J. Acoust. Soc. Am. **65**, 166–178.

Tyler, R. S., and Moore, B. C. J. (**1992**). ''Consonant recognition by some of the better cochlear implant patients,'' J. Acoust. Soc. Am. **92**, 3068–3077.

van den Honert, C., and Stypulkowski, P. H. (**1987**). ''Temporal response patterns of single auditory nerve fibers elicited by periodic electrical stimuli,'' Hearing Res. **29**, 207–222.

van Wieringen, A., and Wouters, J. (**1999**). ''Natural vowel and consonant recognition by LAURA cochlear implantees,'' Ear Hear. (accepted for publication).

Westerman, L. A., and Smith, R. L. (**1984**). ''Rapid and short-term adaptation in auditory nerve responses,'' Hearing Res. **15**, 249–260.

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., and Rabinowitz, W. M. (**1991**). ''Better speech recognition with cochlear implants,'' Nature (London) **352**, 236–238.

Wilson, B. S., Lawson, D. T., Zerbi, M., and Finley, C. C. (**1995**). ''Speech processors for auditory prostheses,'' Final Report, NIH project N01-DC-2-2401, Neural Prosthesis Program, National Institutes of Health, Bethesda, MD.

Wouters, J., Damman, W., and Bosman, A. J. (**1994**). ''Vlaamse opname van woordenlijsten voor spraakaudiometrie,'' Logopedie **7**(6), 28–33.

Mechanisms of fine-surface-texture discrimination in human tactile sensation

Tetsu Miyaoka[a)]
Department of Computer Science, Faculty of Science and Technology, Shizuoka Institute of Science and Technology, 2200-2 Toyosawa, Fukuroi 437-0032, Japan and Department of Autonomic Neuroscience, Division of Higher Nervous Control, Research Institute of Environmental Medicine, Nagoya 464-0814, Japan

Tadaaki Mano
Department of Autonomic Neuroscience, Division of Higher Nervous Control, Research Institute of Environmental Medicine, Nagoya 464-0814, Japan

Masahiro Ohka
Department of Mechanical Engineering, Faculty of Science and Technology, Shizuoka Institute of Science and Technology, 2200-2 Toyosawa, Fukuroi 437-0032, Japan

(Received 10 March 1998; revised 20 August 1998; accepted 11 December 1998)

The purpose of this study was to evaluate the ability of touch to discriminate fine-surface textures and to suggest possible mechanisms of the discriminations. Two experiments were performed. In experiment 1, aluminum-oxide abrasive papers were adopted as stimuli, and psychometric functions and difference thresholds were determined in fine-surface-texture discrimination tasks. The grit values of abrasive papers were 400, 600, 1200, 2000, 3000, 4000, and 8000; corresponding average particle sizes were 40, 30, 12, 9, 5, 3, and 1 μm, respectively. Ten subjects participated in experiment 1. The difference thresholds obtained in experiment 1 were between 2.4 and 3.3 μm. In experiment 2, the tasks were discriminations of ridge height. The cross sections of the etched ridges were rectangular and the ridge heights were 6.3, 7.0, 8.6, 10.8, 12.3, 18.5, and 25.0 μm. Six subjects participated in experiment 2. The difference thresholds in experiment 2 were between 0.95 and 2.0 μm. It was reasoned, based on the Weber fraction values calculated from the difference thresholds and on the limit of neural information-processing ability of humans, that the subjects discriminate fine roughness only from the amplitude information presented in surface unevenness. © *1999 Acoustical Society of America.* [S0001-4966(99)05903-2]

PACS numbers: 43.66.Wv [RVS]

INTRODUCTION

There are many studies which have investigated the perception of surface textures through the sense of touch. Most of them treated the perception of rather rough surfaces: sizes of surface elements were larger than several hundred micrometers. Only a small number of studies have focused on the tactile perception of fine-surface textures: sizes of elements were smaller than 50 μm. However, humans are very good at discriminations of fine-surface textures. Sometimes they discriminate fine-surface textures better by touch than by vision. For example, they can discriminate roughness of abrasive papers with 5-μm particles from 9-μm particles by touch, but cannot discriminate them by vision. This study was designed to investigate possible mechanisms of fine-surface-texture discrimination.

Some previous studies used abrasive papers as stimuli for research of rough-texture perception, but many recent studies have adopted gratings (Goodwin and John, 1991; Johnson and Phillips, 1981; Lederman, 1974; Lederman *et al.*, 1982; Morley *et al.*, 1983; Sathian *et al.*, 1989; Sathian and Zangaladze, 1996; Taylor and Lederman, 1975; Van Boven and Johnson, 1994) or raised dots (Blake *et al.*, 1997; Connor *et al.*, 1990; Connor and Johnson, 1992; Johnson and Hsiao, 1994; Johnson and Lamb, 1981; Lamb, 1983; Lederman *et al.*, 1986) as stimuli. In grating experiments, the researchers found that roughness magnitude increased as grooves became wider (Goodwin and John, 1991; Lederman, 1974; Sathian *et al.*, 1989; Taylor and Lederman, 1975). Goodwin and John (1991) investigated the skin mechanoreceptors which took part in the roughness perception of gratings, and showed that the responses of all three mechanoreceptors (RA, PC, and SA) contributed to the perception of roughness. In raised-dot experiments, Connor *et al.* (1990) found that subjective roughness magnitude was an inverted U-shaped function of dot spacing that peaked near 3-mm spacing. Connor *et al.* (1990) and Blake *et al.* (1997) showed that perceived roughness decreased as dot diameter increased. Johnson *et al.* (Blake *et al.*, 1997; Connor *et al.*, 1990; Connor and Johnson, 1992; Johnson and Hsiao, 1992; Johnson and Hsiao, 1994) had investigated relations between firing of mechanoreceptors of monkeys and perceived roughness of human subjects using embossed-dot patterns as stimuli. They showed that spatial variation in SA (SAI) firing rates had very high correlations with the perceived roughness. Phillips *et al.* (1990) investigated the human mechanoreceptors which were responsible for Braille recognition, and found that SAI discharges showed the most precise resemblance to the Braille pattern.

[a)]All correspondence should be addressed to Tetsu Miyaoka, Department of Computer Science, Faculty of Science and Technology, Shizuoka Institute of Science and Technology, 2200-2 Tovosawa, Fukuroi 437-0032, Japan.

In the small number of studies related to the fine-surface-texture perception of touch, Heller (1989) compared vision and touch in the accuracy of extreme fine-surface-texture discrimination tasks using Japanese abrasive sharpening stones as stimuli. He found that touch was superior to vision for discrimination tasks. Miyaoka and Mano (1991) measured the absolute-detection threshold of fine-surface texture with very fine abrasive papers, and found that the detection threshold was between 1 and 3 μm in particle sizes. Several researchers studied tactile-detection thresholds of edges and dots, not fine textures, with minimal heights. Johansson and LaMotte (1983) measured height-detection thresholds using edges and raised dots as stimuli, and found that the thresholds were 0.85 μm for edges and from 1.09 to 5.97 μm for dots. The detection thresholds of raised dots decreased as the dot diameters increased: the threshold was 5.97, 2.94, and 1.09 μm for dots with diameters of 40, 231, and 602 μm, respectively. LaMotte and Whitehouse (1986) showed that the detection threshold of a raised dot of 550-μm diameter was 2.1 μm, and that the responses of RA alone accounted for the sensory capacity to detect the dot. LaMotte and Srinivasan (1991) measured the detection thresholds of parallel bars and of a matrix of dots. They found the heights of bars and dots at the thresholds were lower than the threshold height of a single dot.

Perception of fine-surface textures by touch is one of the most important functions of tactile information processing. However, there has been only a small number of studies about the perception of fine-surface textures. Also, perceptual mechanisms of fine textures have not been systematically studied. This study was designed to investigate possible mechanisms of fine surface-texture discrimination. There were two purposes for this study. One was to measure the discrimination ability of fine-surface textures. The difference thresholds of fine surface textures were measured in experiment 1. The other was to propose a hypothesis of tactile mechanisms which make the discriminations of fine-surface textures possible. We proposed the hypothesis that subjects discriminate fine-surface textures using amplitude information of the surface unevenness of stimuli. Experiment 2 was performed to measure difference thresholds of ridge-height discrimination, and to check whether subjects had the ability for amplitude discriminations. The results of experiment 2 strengthened the possibility of the hypothesis.

I. EXPERIMENT 1

Experiment 1 adopted fine abrasive papers as stimuli and measured the subjects' discrimination ability of fine-surface textures with the two-alternative, forced-choice technique. There were two reasons for the adoption of abrasive papers or natural stimuli. One reason was that, in the first stage of the study, adopting natural stimuli was a convenient choice to find the characteristics making fine surface-texture discrimination possible. The natural stimuli have many characteristics which might be related to fine-surface discrimination. This study should reveal which characteristic is the most essential for discrimination. The other reason was that producing many different types of stimuli was difficult because of the limits in production techniques. At the beginning of the study, we had no information as to what characteristics might be responsible for the fine-surface discrimination, and could not produce suitable artificial stimuli for the investigation.

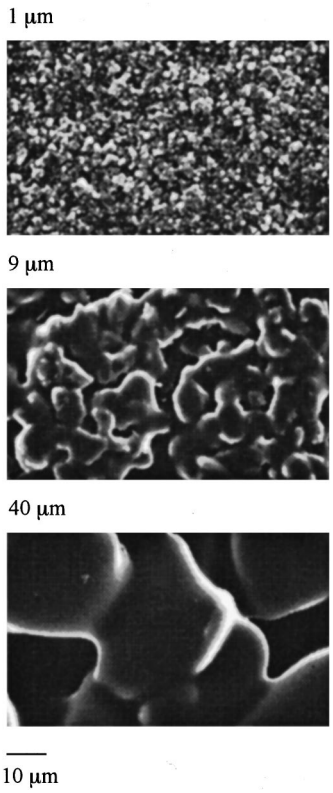


FIG. 1. Pictures of the abrasive-paper surfaces taken with a scanning electron microscope (JEOL, JSM5400). The particle sizes of the abrasive papers are 1, 9, and 40 μm from top to bottom, respectively. The horizontal bar shows a 10-μm length.

A. Method

1. Subjects

Ten subjects, nine men and one woman, aged from 19 to 41 years, took part in experiment 1. Eight of the male subjects were undergraduate students with no previous experience in psychophysical experiments. The other two subjects had previously participated in several other psychophysical experiments.

2. Stimuli

The stimuli were seven aluminum-oxide abrasive papers (Sumitomo 3-M). The grit values assigned by the manufacturer were 400, 600, 1200, 2000, 3000, 4000, and 8000, representing corresponding average particle sizes of 40, 30, 12, 9, 5, 3, and 1 μm, respectively. Observation with a scanning electron microscope (JEOL, JSM5400) confirmed that the sizes of the large particles in each paper corresponded to the particle-size values reported by the manufacturer (Fig. 1). A contact-profile meter (Kosaka, SE-30D) showed that the peak-to-peak amplitudes of the surface unevenness of abrasive papers did not always coincide with the particle-size values of the papers. However, the amplitudes displayed a monotonic relation to the particle size values (Fig. 2).

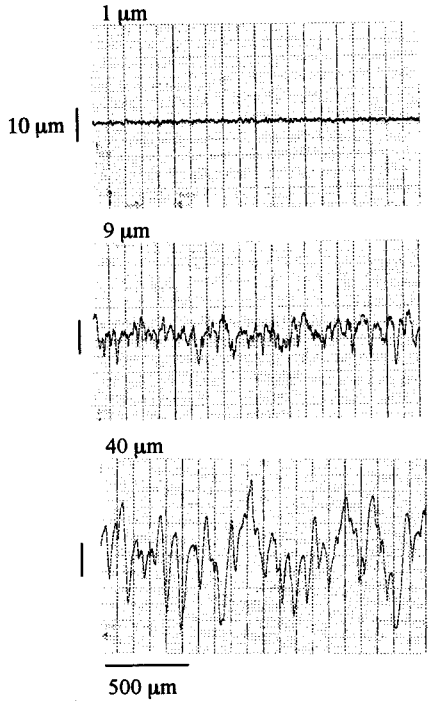


FIG. 2. Cross-section profiles of abrasive papers observed with a contact profile meter (Kosaka, SE-30D). The particle sizes of the papers are 1, 9, and 40 μm from top to bottom, respectively. The three vertical bars show a 10-μm length and the horizontal bar shows a 500-μm length.

The abrasive papers were cut into 5×5-cm squares, and two squares each were glued on boards sized 12×18 cm. Fifty-six stimulus boards were used in experiment 1. Pairwise combinations of seven distinct stimuli plus self-combinations equaled 28 boards. As it was necessary to counterbalance the right and left positions of the squares on each board, the total number of boards was 56.

3. Procedure

The subject was seated in a chair and wore a mask to prevent visual inspection of the test materials. He/she touched the stimuli with the index finger or the middle finger preferred, and determined which abrasive paper on the board felt rougher with the two-alternative, forced-choice technique. Maximum allowed time for the determination was 15 s for each stimulus board. The interstimulus interval was 20 s. Each subject executed ten discrimination trials for each board. The total number of the experimental trials was 560 for each subject. During the experiment, the temperature of the laboratory was maintained between 25 and 27 deg centigrade, and the temperature of the stimulated skin of subjects was greater than 30 deg. Several researchers revealed no or only slight improvements to vibrotactile thresholds when the skin temperature was higher than 30 deg (Bolanowski and Verrillo, 1982; Gescheider *et al.*, 1997; Verrillo and Bolanowski, 1986).

B. Results

Because the experimental results of ten subjects were similar to each other, data was accumulated for each combination, psychometric functions were then calculated. The

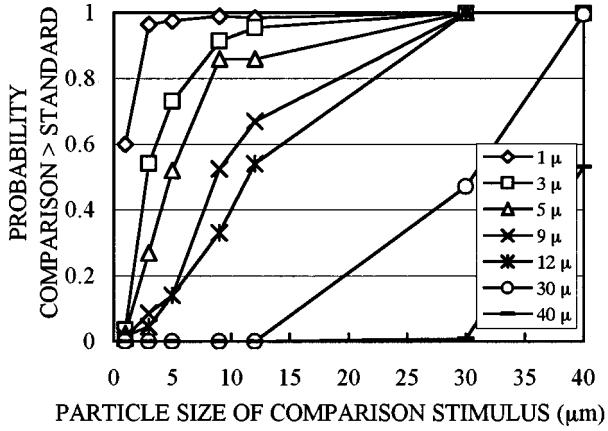


FIG. 3. Psychometric functions based on the data of the fine-surface-texture discrimination experiment. The vertical axis shows the probability that the comparison stimuli were judged rougher than the standard stimuli. The horizontal axis shows the particle sizes of the comparison stimuli. Each symbol in the figure represents the results of each standard stimulus, respectively. The unit of the values in the legend is micrometers.

psychometric functions based on the total data are shown in Fig. 3. Each psychometric function illustrates the probability that the subject judged a comparison stimulus to be rougher than a standard stimulus. For example, in the uppermost function, the subjects judged the 3-μm comparison stimulus to be rougher than the 1-μm standard stimulus with a probability of greater than 95%.

Difference thresholds were determined for four psychometric functions with 3-, 5-, 9-, and 12-μm standard stimuli. It was impossible to determine difference thresholds when the particle sizes of standard stimuli were 1, 30, and 40 μm because these stimuli were almost completely discriminated from other stimuli. The data of the four psychometric functions were normalized and normal-distribution curves were fitted to them with the method of least squares. The difference thresholds were calculated from the fitted curves. They were 2.4, 2.5, 2.7, and 3.3 μm for the particle sizes of 3, 5, 9, and 12 μm, respectively. Weber fractions were calculated as the ratios of the difference thresholds and the particle sizes of corresponding standard stimuli. The values of Weber fractions were 0.81, 0.50, 0.30, and 0.28 for the particle sizes of 3, 5, 9, and 12 μm, respectively (Fig. 4). The fraction values

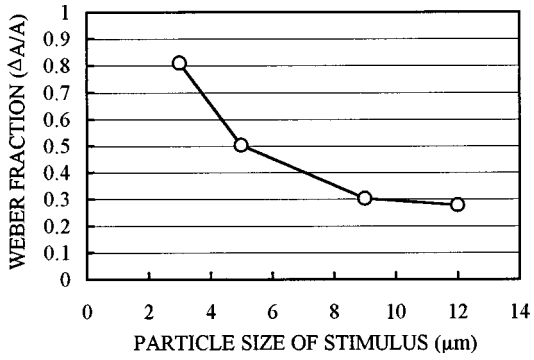


FIG. 4. Weber fractions calculated from the fine-surface-texture discrimination data. The vertical axis shows the Weber fractions and the horizontal axis shows the particle sizes of the stimuli.

changed as a function of particle size, decreasing as the particle size became larger.

C. Discussion

The results of experiment 1 show that human subjects can discriminate fine-surface textures very well with the sense of touch. What tactile information-processing mechanisms make possible the discrimination of fine-surface texture? We propose the hypothesis that subjects discriminate fine-surface textures by the amplitude of the surface unevenness of stimuli. Before the explanation of the ''amplitude information hypothesis,'' we will examine three alternative hypotheses and show that they do not hold.

The first hypothesis is that the subjects use temporal information of surface unevenness for fine-surface discrimination. When they discriminate fine-surface textures, they move their fingers tangentially across the surface to get texture information. Viewing from one point on the stimulated skin, the stimulus unevenness is presented as a kind of vibration, and the information of the vibrotactile stimulus is encoded into neural discharges. Assume that the size of large particles in the stimulus, reported by the manufacturer, represents the vibration wavelength and that the wave form is sinusoidal. For example, the 3-μm stimulus is approximated as a sinusoidal wave with a 3-μm wave length. During experiment 1, it was observed that the subjects moved their fingers at velocities from about 20 to 100 mm/s and gave judgments. If they move their fingers at a velocity of 30 mm/s on the 3- and 30-μm stimuli, they will feel vibrations on the skin equivalent to 10 kHz and 1000 Hz, respectively. On the other hand, mechanoreceptors in the glabrous skin cannot discharge higher than 1000 Hz. If the subjects attempt the discrimination tasks depending on the temporal information of discharge, it is impossible for them to discriminate the 3-μm stimulus from the 30-μm stimulus.

The second hypothesis is that the subjects use spatial information of surface unevenness for the fine-surface-texture discrimination. The density of mechanoreceptive units in the finger tips is lower than 250/cm^2 (Johansson and Vallbo, 1979), and the centers of each receptive field are separated by at least 600 μm. It is obvious that the densities are too low to discriminate the 3-μm from the 30-μm stimulus. Even if each mechanoreceptive unit has several mechanoreceptors (Johansson, 1978; Zelená, 1994), the situation remains the same.

The third hypothesis is that the subjects discriminate surface textures using the friction information between skin and stimuli. A preliminary experiment revealed that no difference existed in the difference-threshold values of fine-surface-discrimination tasks between a low-friction and a high-friction condition. The low-friction condition used stimuli applied with silicone oil (Shin-Etsu Silicone, KF-96-10CS), and the high-friction condition used stimuli without the oil. Although the friction values of the two conditions were different from each other, no difference was found in the results.

Since none of the three alternative hypotheses adequately explains our results, we now consider the hypothesis that the subjects use the amplitude information of surface unevenness for the fine-surface-texture discrimination tasks. We will show that it is possible to discriminate fine-surface textures if the subjects use the amplitude information of the stimulus unevenness. We do this by comparing the results of experiment 1 with the results of a vibrotactile experiment. The Weber fraction was 0.81 for the 3-μm stimulus, and 0.28 for the 12-μm stimulus in experiment 1. The Weber fraction function decreased monotonically as the size of particles increased (Fig. 4). In many sensory modalities, it is observed that the Weber fraction tends to greatly increase at extremely low intensities of stimulus (Gescheider, 1985). As mentioned in the introduction, Miyaoka and Mano (1991) found that the absolute threshold of fine-surface-texture recognition was between 1 and 3 μm. The results of experiment 1 and of Miyaoka and Mano (1991) show that the intensity of the 3-μm stimulus was less than 10 dB SL. Gescheider *et al.* (1990) measured difference thresholds for vibrotactile amplitudes with three methods between 5 and 40 dB. They found that the Weber fractions were from 0.25 to 0.5 at 5 dB and were from 0.15 to 0.3 at 20 dB, depending on the experimental methods. The values of the Weber fractions were independent of stimulus frequency (25 and 200 Hz). Comparison of the Weber fractions of experiment 1 with those of Gescheider *et al.* (1990) showed that the Weber fractions of experiment 1 were larger than the fractions of Gescheider *et al.* at about the same intensity level. These results suggest that it is possible to accept the hypothesis that the subjects performed the fine-surface-texture discrimination tasks in experiment 1 using amplitude differences of surface unevenness of stimulus.

The tactile information-processing system has low-pass filter characteristics and does not pass frequencies higher than 1000 Hz. The tactile system uses the amplitude information passed through the filter to perform the discrimination tasks. Measurements with the contact profile meter (Kosaka, SE-30D) showed that the amplitude information of surface unevenness remained after the wave passed through the low-pass filter. The monotonic relation was maintained between the peak-to-peak amplitudes of surface unevenness and the particle sizes after passing the waves through the filter, which dropped the waves shorter than a 250-μm wavelength (Fig. 5). If the subjects move their fingers on the stimuli at the speed of 30 mm/s, a wave with a 250-μm wavelength has a frequency of 120 Hz. The subjects obtain amplitude information of surface unevenness from waves passed through the filter, and discriminate the fine textures. The results presented here strongly suggest that peak-to-peak amplitude information of surface unevenness, passed through the filter, is the primary source of fine-surface discrimination with touch.

II. EXPERIMENT 2

In experiment 1, the difference thresholds for fine-surface-texture discrimination using abrasive papers as stimuli were measured. The hypothesis was proffered that subjects performed the fine-surface-texture discrimination tasks using the amplitude information of the stimulus surface unevenness. If the subjects really used the amplitude information in the discrimination of fine roughness, they would

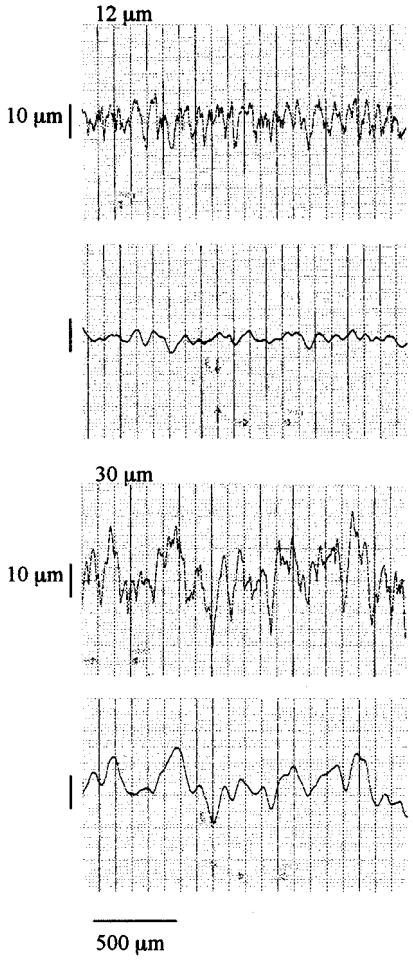


FIG. 5. Cross-section profiles of abrasive papers observed with the contact profile meter with and without a filter. The top two profiles are the cross sections of 12-μm particle abrasive paper. These two profiles show the same part of the same abrasive paper. The upper of the two shows the profile without a filter. The lower shows the profile passed through a filter. The filter passed waves with wavelengths which were greater than 250 μm. The bottom two profiles are the cross sections of 30-μm abrasive paper. The upper shows the profile without a filter. The lower shows the profile passed through the same filter used in the 12-μm abrasive paper.

show at least the same discrimination ability for ridge-height discrimination tasks. In experiment 2, ridge-height discrimination was measured using etched ridges. The purpose of experiment 2 was to measure difference thresholds of ridge heights and to show that the experiment of ridge-height discrimination would have the same or smaller difference thresholds than the abrasive-paper experiment. If the smaller thresholds were found in experiment 2, it would strengthen the possibility of the hypothesis that the subjects use amplitude information for the discrimination of fine-surface texture.

A. Method

1. Subjects

Four males and two females ranging between 20 and 40 years of age participated in the experiment. All subjects except one male were undergraduate students. All of the subjects had experience in psychophysical experiments.

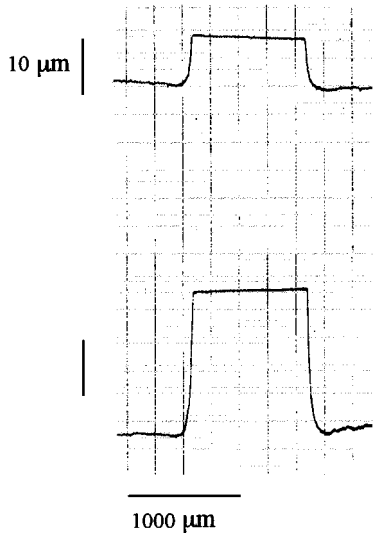


FIG. 6. Cross-section profiles of the ridge stimuli observed with the contact profile meter. The vertical bars show a 10-μm length. The horizontal bar shows 1000 μm. The height of the top ridge is a 8.6 μm and the height of the bottom one is 25 μm.

2. Stimuli and apparatus

The stimuli, ridges 20 mm in length and 1 mm in width, were produced by etching 20×20 mm square stainless-steel plates (SUS 304). The amplitude of surface unevenness on the plate was 0.1 μm before the etching, and was about 0.5 μm after etching, except for the ridge portion, which remained in its original condition. Seven ridges were produced, having heights of 6.3, 7.0, 8.6, 10.8, 12.3, 18.5, and 25.0 μm. Observation with the contact profile meter (Kosaka, SE-30D) showed that the cross sections of the ridges were rectangular (Fig. 6).

During the experiment, the temperature of the glabrous skin of the hand and of the ridge stimuli was controlled. The control device consisted of two parts; one part controlled the temperature of the hand, the other controlled the temperature of the stimuli. The temperature control device was connected to a water circulator (Yamato, CTE82W), which could control the temperature of the device at any temperature between 5 and 45 deg centigrade.

3. Procedure

The subject was seated in a chair and wore an eye mask to prevent visual inspection of the stimuli. He/she touched the stimuli with the index finger or the middle finger of choice and determined which ridge felt higher by means of the two-alternative, forced-choice technique. The maximum time allowed for determination was 20 s for each combination. The interstimulus interval was 20 s. The number of stimulus combinations was 56, determined by the same procedure as in experiment 1. Each subject performed ten trials for each combination of the ridge stimuli; therefore, the number of total experimental trials for each subject was 560. The temperature of the stimuli and the stimulated area of the skin was maintained at 30±1 deg centigrade during the experiment.

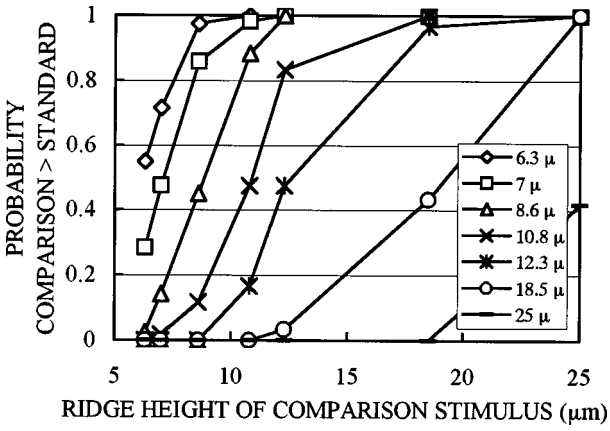


FIG. 7. Psychometric functions based on the data of the ridge-height discrimination experiment. The vertical axis shows the probability that comparison stimuli were judged higher than standard stimuli. The horizontal axis shows the ridge heights of the comparison stimuli. Each symbol in the figure shows the results of each standard stimulus, respectively. The unit of the values in the legend is micrometers.

B. Results

Data were accumulated for all subjects for each combination, psychometric functions were then calculated. The psychometric functions of ridge-height discrimination are shown in Fig. 7. Each psychometric function shows the probability of the comparison stimulus being judged higher than the standard stimulus. For example, when the height of the standard stimulus was 7.0 μm (open squares), the 8.6-μm comparison stimulus was judged higher than the standard stimulus with a probability of 0.86.

The difference thresholds of the ridge-height discrimination tasks were calculated from the curves fitted to the data of the psychometric function by the method of least squares. The four difference thresholds were 0.95, 1.1, 1.2, and 2.0 μm for the standard stimuli of 7.0, 8.6, 10.8, and 12.3 μm, respectively. The difference thresholds were not determined for the 6.3-, 18.5-, and 25.0-μm standard stimuli because they were almost perfectly discriminated from all other stimuli. Weber fractions were calculated as the ratios of corresponding difference thresholds and the ridge heights. The Weber fractions were 0.14, 0.13, 0.11, and 0.16 for 7.0-, 8.6-, 10.8-, and 12.3-μm ridges, respectively (Fig. 8).

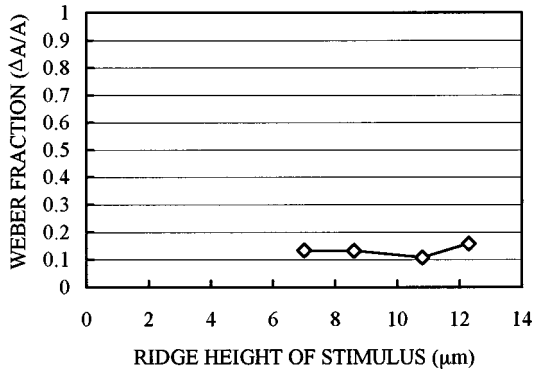


FIG. 8. Weber fractions calculated from the ridge-height discrimination data. The vertical axis shows the Weber fractions and the horizontal axis shows the ridge heights of the stimuli.

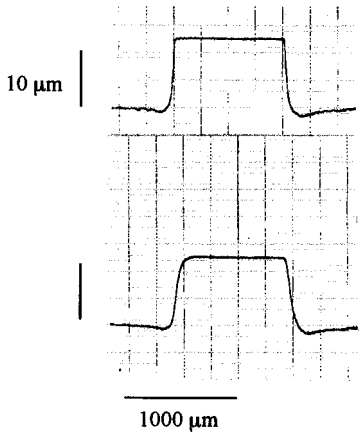


FIG. 9. Cross-section profiles of a ridge observed with the contact profile meter with and without a filter. The top is the profile of the 12.3-μm ridge without a filter. The bottom is the profile, passed through a filter, of the same part of the same ridge shown on the top. The filter passed waves with wavelengths greater than 250 μm.

C. Discussion

Experiment 2 was performed to measure the difference thresholds and to strengthen the possibility of the hypothesis that humans discriminate fine-surface textures using amplitude information of stimulus surface unevenness. The difference thresholds of ridge-height discrimination were from 0.95 to 2.0 μm. These results showed that humans can discriminate very small differences in ridge heights using tactile sensation. The Weber fractions, calculated from the difference thresholds, were 0.14, 0.13, 0.11, and 0.16 for 7.0-, 8.6-, 10.8-, and 12.3-μm ridges, respectively. The Weber fractions showed no increasing nor decreasing tendency in their values depending on the changing of ridge heights (Fig. 8). This phenomenon means that the absolute-detection threshold of ridge height is considerably lower than 6 μm. Johansson and LaMotte (1983) determined that the detection threshold of edge height was 0.85 μm, and LaMotte and Whitehouse (1986) found that the detection threshold of a raised dot, with a diameter of 550 μm, was 2.1 μm. Assuming that the detection threshold of a ridge is 1 μm, then the 7.0-, 8.6-, 10.8-, and 12.3-μm ridges are 17, 19, 21, and 22 dB SL, respectively. Comparing the Weber fractions of ridge heights with the fractions of a vibrotactile stimuli (Gescheider *et al.*, 1990) at the same intensity level, the Weber fractions of the ridge heights were the same or only slightly smaller than those of vibrotactile stimuli. The results of the ridge-height experiment support the hypothesis that the subjects used amplitude information of surface unevenness in the discrimination of ridge height. Tactile nervous systems have low-pass filter characteristics. The ridge height, however, remained the same after passing through the filter, which excluded wavelengths shorter than 250 μm, as shown in Fig. 9. It is logical to assume that the subjects perceived the ridge heights as waves because they moved their fingers tangentially across the ridge stimuli in the experiment.

Comparing the Weber fractions of experiment 2 with the fractions of the fine-surface-texture discrimination experiment (experiment 1), the former fractions were smaller than the latter ones. The Weber fractions of experiment 2 were

from 0.11 to 0.16. On the other hand, the fractions of experiment 1 were from 0.28 to 0.81. When comparing the Weber fractions of the texture experiment and the ridge experiment with similar-sized standard stimuli, the Weber fraction of the texture experiment was about two times larger than that of the ridge experiment; the Weber fractions were 0.28 for texture and 0.16 for ridges when the standard stimulus sizes were 12 μm for the texture and 12.3 μm for the ridge, respectively. A possible reason for this result is as follows: during the ridge experiment, the discrimination tasks were easy because the stimulated mechanoreceptors discharged simultaneously at the ridge; however, during the texture experiment the tasks were more difficult because the mechanoreceptors do not show these simultaneous discharges.

The stimuli in experiment 2 were simple, and the results strongly suggested that the subjects used amplitude information for the discrimination of ridge heights. If the subjects discriminate the ridge heights using the amplitude information of ridges which passed through a low-pass filter, they can use the same ability for the discrimination of fine-surface textures. The results of experiment 2 strengthened the possibility of the hypothesis that the subjects discriminate fine-surface textures depending on the amplitude information of surface unevenness.

III. CONCLUSION

The purpose of this study was to investigate possible tactile mechanisms for making fine-surface-texture discriminations. In experiment 1, the discrimination of fine-surface textures at the finger tips, using abrasive papers as stimuli, was measured. The Weber fractions of texture discriminations were between 0.28 and 0.81. Based on the results of experiment 1 and other psychophysical and neurophysiological data, the following hypothesis was proposed: the subjects perform the discrimination tasks based on the amplitude information of surface unevenness.

If the amplitude information hypothesis is valid, the subjects would have shown at least the same ability in ridge-height discriminations as in texture discriminations. In experiment 2, the difference thresholds for the discrimination of ridge height was measured. The Weber fractions calculated from the difference thresholds were between 0.11 and 0.16. The results of experiment 2 strengthen the possibility of the hypothesis. Further experimentation is needed to determine if these explanations can account for the results reported here.

ACKNOWLEDGMENTS

This work was supported by a Grant-in-Aid for Scientific Research (C) from the Ministry of Education, Science and Culture, No. 07610093, and by the Mikiya Science and Technology Foundation.

Blake, D. T., Hsiao, S. S., and Johnson, K. O. (**1997**). ''Neural coding mechanisms in tactile pattern recognition: the relative contributions of slowly and rapidly adapting mechanoreceptors to perceived roughness,'' J. Neurosci. **17**, 7480–7489.

Bolanowski, S. J., Jr., and Verrillo, R. T. (**1982**). ''Temperature and criterion effects in a somatosensory subsystem: a neurophysiological and psychophysical study,'' J. Neurophysiol. **48**, 836–855.

Connor, C. E., Hsiao, S. S., Phillips, J. R., and Johnson, K. O. (**1990**). ''Tactile roughness: neural codes that account for psychophysical magnitude estimates,'' J. Neurosci. **10**, 3823–3836.

Connor, C. E., and Johnson, K. O. (**1992**). ''Neural coding of tactile texture: comparison of spatial and temporal mechanisms for roughness perception,'' J. Neurosci. **12**, 3414–3426.

Gescheider, G. A. (**1985**). *Psychophysics: Method, Theory, and Application* (Erlbaum, Hillsdale, NJ), Chap. 1, pp. 1–36.

Gescheider, G. A., Bolanowski, S. J., Jr., Verrillo, R. T., Arpajian, D. J., and Ryan, T. F. (**1990**). ''Vibrotactile intensity discrimination measured by three methods,'' J. Acoust. Soc. Am. **87**, 330–338.

Gescheider, G. A., Thorpe, J. M., Goodarz, J., and Bolanowski, S. J. (**1997**). ''The effects of skin temperature on the detection and discrimination of tactile stimulation,'' Somato. Motor Res. **14**, 181–188.

Goodwin, A. W., and John, K. T. (**1991**). ''Peripheral neural basis for the tactile perception of texture,'' in *Information Processing in the Somatosensory System*, edited by O. Franzén and J. Westman (Stockton, New York), Chap. 7, pp. 81–91.

Heller, M. A. (**1989**). ''Texture perception in sighted and blind observers,'' Percept. Psychophys. **45**, 49–54.

Johansson, R. S. (**1978**). ''Tactile sensibility in the human hand: receptive field characteristics of mechanoreceptive units in the glabrous skin area,'' J. Physiol. (London) **281**, 101–123.

Johansson, R. S., and LaMotte, R. H. (**1983**). ''Tactile detection thresholds for a single asperity on an otherwise smooth surface,'' Somato. Res. **1**, 21–31.

Johansson, R. S., and Vallbo, Å. B. (**1979**). ''Tactile sensibility in the human hand: Relative and absolute densities of four types of mechanoreceptive units in glabrous skin,'' J. Physiol. (London) **286**, 283–300.

Johnson, K. O., and Hsiao, S. S. (**1992**). ''Neural mechanisms of tactual form and texture perception,'' Annu. Rev. Neurosci. **15**, 227–250.

Johnson, K. O., and Hsiao, S. S. (**1994**). ''Evaluation of the relative roles of slowly and rapidly adapting afferent fibers in roughness perception,'' Can. J. Physiol. Pharmacol **72**, 488–497.

Johnson, K. O., and Lamb, G. D. (**1981**). ''Neural mechanisms of spatial tactile discrimination: neural patterns evoked by Braille-like dot patterns in the monkey,'' J. Physiol. (London) **310**, 117–144.

Johnson, K. O., and Phillips, J. R. (**1981**). ''Tactile spatial resolution. I. Two-point discrimination, gap detection, grating resolution, and letter recognition,'' J. Neurophysiol. **46**, 1177–1191.

Lamb, G. D. (**1983**). ''Tactile discrimination of textured surfaces: peripheral neural coding in the monkey,'' J. Physiol. (London) **338**, 567–587.

LaMotte, R. H., and Srinivasan, M. A. (**1991**). ''Surface microgeometry: tactile perception and neural encoding,'' in *Information Processing in the Somatosensory System*, edited by O. Franzén and J. Westman (Stockton, New York), Chap. 4, pp. 49–58.

LaMotte, R. H., and Whitehouse, J. (**1986**). ''Tactile detection of a dot on a smooth surface: peripheral neural events,'' J. Neurophysiol. **56**, 1109–1128.

Lederman, S. J. (**1974**). ''Tactile roughness of grooved surfaces: the touching process and effects of macro- and microsurface structure,'' Percept. Psychophys. **16**, 385–395.

Lederman, S. J., Loomis, J. M., and Williams, D. A. (**1982**). ''The role of vibration in the tactual perception of roughness,'' Percept. Psychophys. **32**, 109–116.

Lederman, S. J., Thorne, G., and Jones, B. (**1986**). ''Perception of texture by vision and touch: multidimentionality and intersensory integration,'' J. Exp. Psychol. **12**, 169–180.

Miyaoka, T., and Mano, T. (**1991**). ''Touch discrimination ability of humans using fine surface textures as stimuli,'' Ann. Res. Inst. Environ. Med. **42**, 249–252.

Morley, J. W., Goodwin, A. W., and Darian-Smith, I. (**1983**). ''Tactile discrimination of gratings,'' Exp. Brain Res. **49**, 291–299.

Phillips, J. R., Johansson, R. S., and Johnson, K. O. (**1990**). ''Representation of Braille characters in human nerve fibers,'' Exp. Brain Res. **81**, 589–592.

Sathian, K., Goodwin, A. W., John, K. T., and Darian-Smith, I. (**1989**). ''Perceived roughness of a grating: correlation with responses of mechanoreceptive afferents innervating the monkey's fingerpad,'' J. Neurosci. **9**, 1273–1279.

Sathian, K., and Zangaladze, A. (**1996**). ''Tactile spatial acuity at the human fingertip and lip: bilateral symmetry and inter-digit variability,'' Neurology **46**, 1464–1466.
Taylor, M. M., and Lederman, S. J. (**1975**). ''Tactile roughness of grooved surfaces: a model and the effect of friction,'' Percept. Psychophys. **17**, 23–36.
Van Boven, R. W., and Johnson, K. O. (**1994**). ''The limit of tactile spatial resolution in humans: grating orientation discrimination at the lip, tongue, and finger,'' Neurology **44**, 2361–2366.
Verrillo, R. T., and Bolanowski, S. J., Jr. (**1986**). ''The effects of skin temperature on the psychophysical responses to vibration on glabrous and hairy skin,'' J. Acoust. Soc. Am. **80**, 528–532.
Zelená, J. (**1994**). *Nerves and Mechanoreceptors* (Chapman & Hall, London), Chap. 5, pp. 186–217.

Target detection by an echolocating harbor porpoise (*Phocoena phocoena*)

R. A. Kastelein
Harderwijk Marine Mammal Park, Strandboulevard Oost 1, 3841 AB Harderwijk, The Netherlands

W. W. L. Au
Hawaii Institute of Marine Biology, University of Hawaii, P.O. Box 1106, Kailua, Hawaii 96734

H. T. Rippe and N. M. Schooneman
Harderwijk Marine Mammal Park, Strandboulevard Oost 1, 3841 AB Harderwijk, The Netherlands

(Received 30 July 1998; revised 6 November 1998; accepted 11 December 1998)

Two echolocation experiments are described. They were conducted on the same harbor porpoise housed in a sea pen, one year apart at Neeltje Jans, The Netherlands. The aims were to determine the target detection ability of an echolocating harbor porpoise, with the ultimate goal to predict the distance at which harbor porpoises can detect fishing nets. In experiment 1, the maximum distance at which the 3-year-old porpoise could detect a 7.62-cm diameter water-filled stainless-steel sphere by echolocation was determined psychophysically. The 50%-correct detection threshold was reached when the sphere was at a distance of 26 m from the porpoise's rostrum. In experiment 2, conducted a year later, the maximum detection distance for a 5.08-cm water-filled stainless-steel sphere was 15.9 m. The target strengths of both targets were measured using simulated harbor porpoise echolocation signals and the results, coupled with transmission-loss calculations, indicated that the echo levels received by the porpoise with the targets at the threshold ranges in the two experiments were only 1.3 dB apart. Together with information on the target strengths of various fishing nets, the results of the present study can be used to predict the distance at which the nets can be detected by harbor porpoises. © *1999 Acoustical Society of America.*
[S0001-4966(99)04803-1]

PACS numbers: 43.80.Ev, 43.80.Jz, 43.80.Lb [FD]

INTRODUCTION

Annually, tens of thousands of harbor porpoises (*Phocoena phocoena*) die in gillnets (Jefferson and Curry, 1994). Why they become entangled is not clear, but the porpoises are possibly unable to detect the nets with echolocation in time to avoid them. Very little is known about the echolocation abilities of porpoises, although the echolocation signals produced by them are known to have a peak frequency of around 130 kHz (Møhl and Andersen, 1973; Kamminga and Wiersma, 1981; Verboom and Kastelein, 1995, 1997) and the hearing-frequency range of harbor porpoises is wide (Andersen, 1970; Bibikov, 1992). Target-detection ranges can be used to estimate the distances at which harbor porpoises are able to detect various fishing nets. The detection abilities of harbor porpoises have only been tested once. Busnell and Dziedzic (1967) conducted an obstacle-avoidance experiment and found that a blindfolded harbor porpoise could detect a 0.2-mm diameter metallic wire, a 0.88-mm diameter perlon thread, and a 1-mm diameter nylon monofilament line at the 50%-detection level with detection distances between 4 and 11 cm.

Harbor porpoises are not kept in captivity permanently, and are therefore not normally available for the psychophysical testing of echolocation performance. However, when a stranded juvenile harbor porpoise, which had been raised by tube feeding, was moved to a large floating pen in a sea harbor, the opportunity arose to conduct echolocation experiments with the ultimate goal to determine the detectability of fishing nets to harbor porpoises. The target-detection range of the echolocating harbor porpoise was determined by testing its ability to detect 5.08- and 7.62-cm diameter water-filled stainless-steel spheres.

I. MATERIALS AND METHODS

The study animal was a stranded male harbor porpoise (code PpSH030) which had been raised at the Netherlands Cetacean Research and Rehabilitation Center at the Harderwijk Marine Mammal Park from the approximate age of 5 weeks, when he weighed 12.9 kg and had a body length of 89 cm. During the course of experiment 1, the animal was 3 years old, his weight increased from 29.5 to 30.9 kg, and his length from 129.0 to 129.5 cm. His girth in front of the pectoral fins (near the ear) increased from 62 to 64.5 cm. During the course of experiment 2, the study animal was 4 years old, his weight increased from 34.4 to 35.8 kg, and his length from 132.5 to 134 cm. His girth in front of the pectoral fins was 67 cm. In both experiments, the study animal was kept with another male harbor porpoise (PpSH036) of similar size and age.

Both experiments were done at the Rehabilitation and Research Center Neeltje Jans in the southwest Netherlands (51°37′N, 03°40′E), which is operated by the Harderwijk Marine Mammal Park and Waterland Neeltje Jans. The animal was housed in a floating pen (34×20 m; 3.5-m deep at the sides, and 5.0-m deep in the center during high tide, and 3.5 m during low tide during experiment 1 (Fig. 1) and 2.4-m

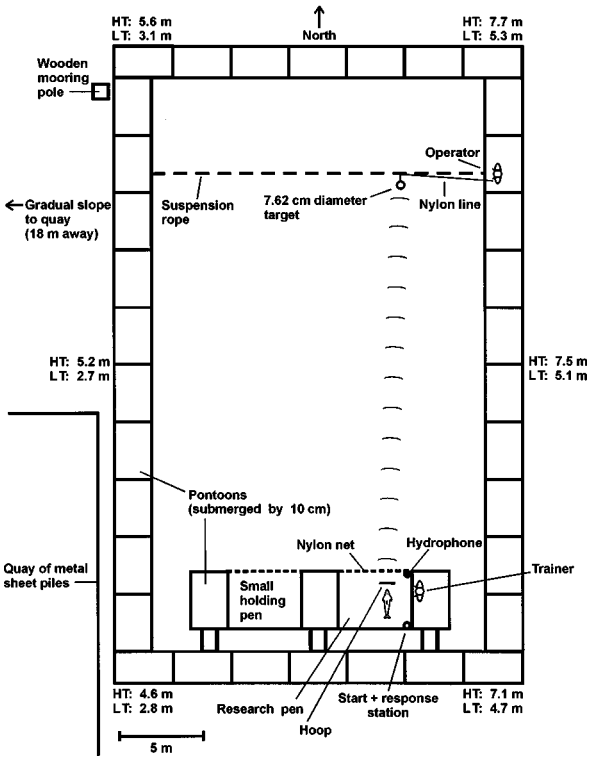


FIG. 1. The study area during experiment 1, showing the floating pen containing the small holding and experimental pens, as well as the station, the hoop, and the target. The water depths at high and low tide are indicated (HT=high tide, LT=low tide).

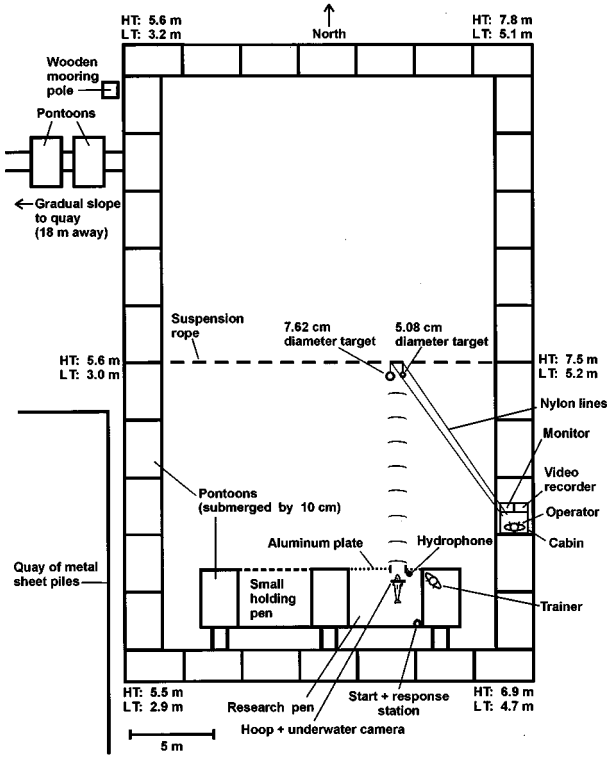


FIG. 2. The study area during experiment 2, showing the floating pen containing the small holding and research pens, as well as the station buoy, the hoop, the target, the equipment, and the operator's cabin. The water depths at high and low tide are also shown.

deep at the sides, and 2.8-m deep in the center during experiment 2; Fig. 2). The surrounding pontoons (plywood boxes filled with Styrofoam and coated with fiberglass) were submerged by 10 cm. The floor and sides of the pen were made of net (twine thickness: 3 mm; stretched mesh size: 9 cm), so the surrounding seawater could flow through it. The twine was covered with algae and bivalve mollusks.

During experiment 1, the salinity was measured weekly and varied between 3.3% and 3.5%, and the mean monthly water temperature was $15.6\pm1.6\,^{\circ}\mathrm{C}$ ($n=30$, range 13–19 °C) in September, and $13.3\pm0.9\,^{\circ}\mathrm{C}$ ($n=28$, range 12–14 °C) in October. The water clarity, measured with a secchi disc, was on average 2.8 m±0.8 m ($n=56$, range 1.1–4.5 m). During experiment 2, the salinity was measured weekly and varied between 3.4% and 3.6%, and the mean monthly water temperature was $12.6\pm1.6\,^{\circ}\mathrm{C}$ ($n=29$, range 9.9–14.9 °C) in May, $15.8\pm0.9\,^{\circ}\mathrm{C}$ ($n=30$, range 14.0–17.4 °C) in June, $18.3\pm1.5\,^{\circ}\mathrm{C}$ ($n=24$, range 16–21 °C) in July, and $21.2\pm1.5\,^{\circ}\mathrm{C}$ ($n=31$, range 18–23 °C) in August. The water clarity, measured with a secchi disc, was on average 3.5 m±0.7 m ($n=78$, range 1.9–4.9 m).

The harbor around the pen was horseshoe-shaped (500 $\times$280 m) with the entrance to the northeast. No shipping occurred within 2 km of the pen during the study. The sea floor below the floating pen was flat and covered with sandy silt. The water depth at the location of the floating pen varied depending on the tide (Figs. 1 and 2). The pen was parallel to the shore, and there were no acoustical obstacles (except possibly fish) within 100 m of the animal, in the direction in which it had to echolocate during the experiments.

Before each session began, the study animal's poolmate was put in a small (3.6$\times$2.9 m; 1.2-m deep) holding pen so that he could not disturb the study animal (Figs. 1 and 2). The poolmate was kept occupied by a trainer to make his stay in the holding pen a positive event. During experiments, the study animal was in a small (3.6$\times$2.9 m; 1.2-m deep) research pen (Figs. 1 and 2).

A. Experiment 1

A 7.62-cm diameter water-filled stainless-steel sphere served as a target, and was hung from a nylon suspension rope which was hung across the width of the pool (Fig. 1). The operator could remove the target from the water or lower it to a depth of 1.0 m by means of a nylon monofilament line. Echolocation clicks were made audible to the researchers by means of a custom-made hydrophone and a heterodyne bat detector (Batbox III, Stag Electronics, UK) which was tuned at 120 kHz (Møhl and Andersen, 1973). Because the animal echolocated shortly before responding, this indicated that the animal used echolocation (and no unforeseen accidental cues) to determine the presence or absence of a target.

A trial began with the target usually out of the water and the animal at the station (a buoy), facing away from the target. The target was slowly lowered by the operator when the animal was stationed at the buoy. The target was submerged in target trials or lowered to just above the water surface in target-absent (control) trials. The operator moved the target as quietly as possible. To minimize the provision of inadvertent audible cues to the animal, the target was

sometimes left in the water between successive target-present trials, and sometimes lowered into the water and lifted again during target-absent trials.

When the trainer rang a bell, the animal swam into the stationing hoop (a 25-cm diameter hoop made of a 2.5-cm diameter plastic tube which was suspended 20 cm above the floor of the research pen), echolocating at the target as it swam. A go/no-go response paradigm was used in which the animal, after echolocating, would swim back to the station to indicate a target-present response or remain in the stationing hoop until a dog whistle was blown to indicate a target-absent response. After each correct target-present and target-absent response, the animal received a reward in the form of a fish (herring, *Clupea harengus* or Norway pout, *Trisopterus esmarkii*). The size of the fish reward was the same for correct target-present and target-absent trials, and sessions were conducted at the usual feeding times, when the study animal had not eaten for 1.5 h. Any leftover fish was fed to the animal after the session.

Each session consisted of 6–16 trials and lasted for about 20 min. Each session consisted of approximately 50% target-present and 50% target-absent trials in pseudorandom order (never more than two consecutive target-present or -absent trials; Zar, 1984).

The trainer communicated with the operator via a two-way radio. Before a trial began, the trainer did not know whether the target would be present or absent. When the porpoise left the station, the trainer told the operator how it responded, and the operator told the trainer whether to reward or not. The operator recorded the animal's responses.

To determine the approximate detection threshold, the distance between the animal (while in the hoop) and the target was increased in 2.5-m increments from 7.5 to 20 m. The animal's performance was above 90% in at least 18 target-present trials (about three sessions) at each distance before the distance was increased. Then, when trials were conducted at a target distance of 30 m, the animal did not detect the sphere. This suggested that the threshold was between 20 and 30 m. Therefore, target distances between 20 and 27.5 m were used in random order (Zar, 1984) between sessions in 2.5-m steps.

Two sessions were held per day, except when it rained, or when the estimated wind speed was over Beaufort 5. Fifty-seven sessions were conducted between 31 August and 28 October 1996.

B. Experiment 2

In addition to the target used in experiment 1, a 5.08-cm diameter water-filled stainless-steel sphere was also used in experiment 2. Each target was suspended from a nylon monofilament fishing line which ran through a noose of a nylon suspension rope which was hung across the width of the pool 95 cm above the water surface (Fig. 2). From inside a cabin, the operator could lower each target to a depth of 1.0 m by means of the nylon line (Fig. 2). Echolocation clicks were made audible to the researchers as in experiment 1.

A trial usually began with both targets out of the water and the animal at the station (a buoy), facing away from the targets (Fig. 2). When the animal was stationed, the trainer signaled to the operator, who then slowly lowered one target to 1.0 m below the water surface in target-present trials, or to just above the water surface in target-absent trials. The operator moved the target as quietly as possible and very slowly to prevent the target from swinging. To minimize inadvertent audible cues to the animal, the target was sometimes left in the water between successive target-present trials, sometimes lowered into the water and lifted again before the start of target-absent trials, and lowered to just above the water in target-absent trials. This was done according to a predetermined random schedule. The operator signaled to the trainer when a target was in position, but did not indicate which target had been used, or which position it was in (double-blind test).

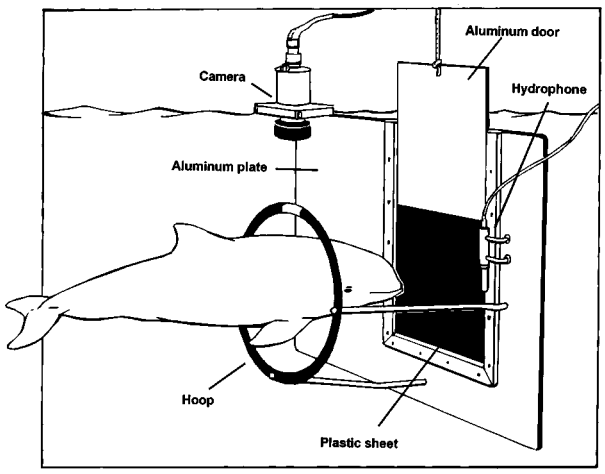


FIG. 3. The study animal in the hoop with the sliding door open during experiment 2 (Drawing: Rijkent Vleeshouwer).

When the trainer rang a bell, the animal swam into a 49-cm diameter hoop which was 40 cm from an aluminum plate (110×128 cm, 3 mm thick; Fig. 3) which acted as an acoustic screen. To dampen sounds, the plate was covered with closed-cell neoprene on the side of the animal. The porpoise was trained to station with his pectoral fins level with the hoop, his blowhole about 70 cm below the water surface, and the tip of his rostrum about 80 cm below the water surface. When he had remained in this position for about 2 s, the trainer opened an aluminum sliding door (45 ×45 cm, 3 mm thick) in the aluminum plate (Fig. 3). The doorway was obscured with black plastic (part of a waste bag) to prevent the porpoise from using vision to detect the target. Tests with 120-kHz sound signals showed that the plastic caused no attenuation.

The go/no-go response paradigm was used. After the animal echolocated, he swam back to the station buoy to indicate a target-present response or remained in the hoop until a whistle was blown (about 6 s after the door had been opened) to indicate a target-absent response. The porpoise was filmed from above with an underwater video camera, and the images were visible to the operator on a monitor in the cabin. The operator would signal to the trainer whether the response was correct or incorrect. For incorrect target-present responses (misses), the porpoise remained in the hoop until the trainer tapped the station buoy (about 6 s after the door was opened). The porpoise was rewarded with a fish

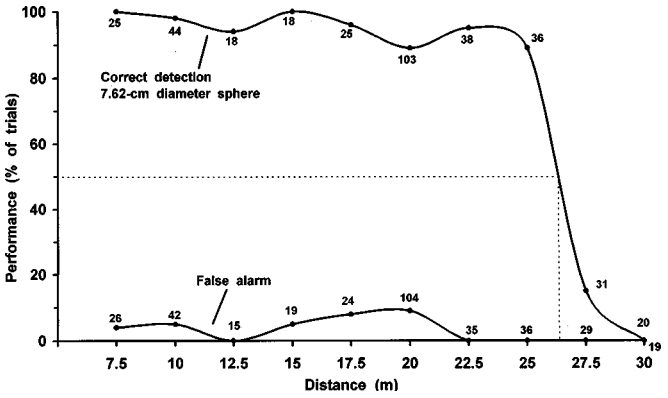


FIG. 4. The percentage of trials in which the 7.62-cm diameter target was correctly detected, and percentage of target-absent trials which resulted in a false alarm, both as a function of distance during experiment 1. The numbers near the lines indicate the sample sizes (number of trials).

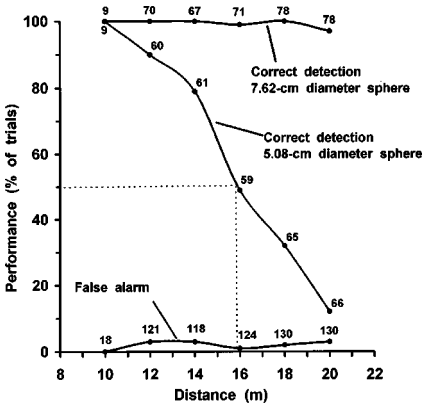


FIG. 5. The percentage of trials in which the 5.08- and 7.62-cm diameter targets were correctly detected, and the percentage of target-absent trials which resulted in a false alarm, both as a function of distance during experiment 2. The numbers near the lines indicate the sample sizes (number of trials).

(sprat, *Sprattus sprattus*) only when it responded correctly. Sessions were conducted at the usual feeding times. The porpoise had not eaten for 1.5 h before each session, and any fish left over was fed to him after the session.

Each session consisted of 12–23 trials and lasted for 10–20 min. Each session consisted of approximately 50% target-present and 50% target-absent trials in pseudorandom order: never more than three present or absent trials were conducted in succession. Random numbers were obtained from the Gellerman series.

Two targets were used in experiment 2. The 7.62-cm diameter target was used to maintain stimulus control. Results of experiment 1 indicated that the animal was able to detect the 7.62-cm diameter target almost perfectly for distances up to 25 m. The 7.62-cm diameter target was presented in 50% of the target-present trials at the same distance as the 5.08-cm diameter target. So, a session consisted of three to six 5.08-cm diameter target-present trials, three to six 7.62-cm diameter target-present trials, and 6 to 12 target-absent trials. The 5.08-cm target was never presented in two successive trials. Sometimes other behaviors were solicited from the porpoise between trials to keep him motivated.

Before each trial began, the trainer did not know whether a target would be present or absent. She was sitting at a 135 deg angle to the porpoise-target axis and did not look towards the target (Fig. 2). When the porpoise left the station, the operator observed the animal's behavior on a monitor in the cabin (Fig. 2), told the trainer whether or not to reward the porpoise, and recorded the animal's responses.

To determine the detection threshold, the target was initially placed 4 m away from the animal's rostrum. When the animal responded correctly in 90% of all trials in three successive sessions, the target was moved 2.0 m further away and the procedure repeated. The approximate threshold was thus determined to be close to 16 m. Thereafter, the target distance for each session was randomly chosen from 12, 14, 16, 18, and 20 m.

Two sessions were held per day, except when it rained, or when the estimated wind speed was over Beaufort 5 (causing waves and large movements of the pen and thus much underwater noise). One-hundred-and-four sessions were conducted between 15 June and 1 September 1997.

The echolocation signal of the porpoise was measured with a B&K 8103 hydrophone and recorded on a Racal-store 7 tape recorder operating at 76.2 cm/s providing a 170-kHz bandwidth. The target strength of the 7.62- and 5.08-cm diameter spheres was measured in a test tank by projecting a replica of the porpoise echolocation signal. A monostatic system using the same transducer for projecting the incident signal and receiving the echo signal was used (Au and Snyder, 1980). The targets were suspended with a thin line approximately 2 m from the transducer. The incident signal was projected after initial movements of the target damped down and the echoes were essentially constant. Target strength is defined as the ratio of the incident signal to the echo signal referenced to a distance of 1 m from the target (Urick, 1983). It was measured in terms of both the peak-to-peak values of the incident and echo signals and in terms of the energy in the incident and echo signals.

II. RESULTS

The porpoise's performance in experiment 1, using the 7.62-cm target, is shown in Fig. 4. The 50%-correct detection threshold was reached when the target was at a distance of 26 m from the porpoise's rostrum. The false-alarm level remained below 10% of trials throughout the experiment. The animal showed motivational problems when the target was presented at a distance of 20 m. This behavior was corrected by using the time-out technique. Therefore, the sample size is larger at this distance than at other distances. During the actual experiment (with the target presented at random distances between 20 and 27.5 m), the false-alarm level was 0%.

The porpoise's performance in experiment 2 is shown in Fig. 5. The percentage of correct detections for the 5.08-cm diameter sphere decreased monotonically as the distance increased. The 50%-detection threshold for the 5.08-cm diameter target was 15.9 m. The animal could detect the 7.62-cm diameter target up to the maximum distance tested (20 m). The false-alarm level during target-absent trials remained below 3% throughout the experiment.

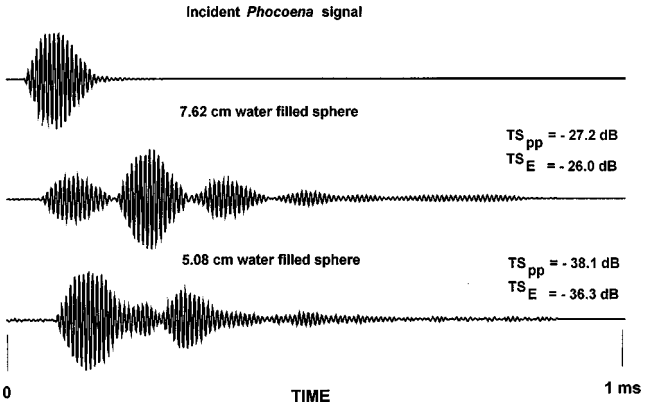


FIG. 6. The waveform of a simulated harbor porpoise signal and the waveforms of the echoes and target strengths of the 5.08- and 7.62-cm diameter water-filled stainless-steel spheres. The target strength based on the incident and reflected energy was calculated for an integration time of 1 ms.

The results of the target strength measurements are shown in Fig. 6. The 7.62-cm diameter sphere had a peak-to-peak target strength of −27.2 dB and a target strength based on energy of −26.0 dB. The 5.08-cm diameter sphere had a peak-to-peak target strength of −38.1 dB and a target strength based on energy of −36.3 dB.

III. DISCUSSION AND CONCLUSIONS

After experiment 1, potential visible or audible cues given by the trainer or the operator, or visible cues from the target were ruled out. In experiment 2 the operator was in a cabin (invisible to the porpoise), and the trainer did not look at the targets during the sessions. The target movements were randomly scheduled in experiment 2. The aluminum plate and plastic sheet behind the sliding door prevented the animal from using visual cues, and the aluminum plate prevented the animal from echolocating the target before the sliding door was opened. Thus, the porpoise was always in the same position when echolocating, so its movements could be filmed.

Two targets were used in experiment 2 to maintain better stimulus control. The sample size per distance was larger in experiment 2 than in experiment 1; more trials per session could be conducted because the animal was more experienced.

From the results of this study, it is possible to determine the echo levels received by the porpoise at the threshold ranges. The porpoise typically emitted echolocation signals with peak frequencies between 125–135 kHz. The ambient noise at those frequencies was extremely low (probably at the thermal limit) and was not measurable with a low-noise hydrophone. Therefore, the noise-limited form of the sonar equation is not applicable, and an estimate of the echo level pertinent to the target-detection range is all that is needed. The auditory system of dolphins and other mammals typically process acoustic information by detecting acoustic energy (Johnson, 1968; Au *et al.*, 1988). Therefore, the energy flux density in the echo is the appropriate parameter to determine. The energy flux density in the echo (EE) in dB can be expressed as

$$\mathrm{EE}=\mathrm{SE}-2\mathrm{TL}+\mathrm{TS}_E, \tag{1}$$

where SE is the source energy flux density, TL is the one-way transmission loss, and TS_E is the target strength based on energy. The transmission loss consists of spherical spreading loss plus a sound-absorption term and can be expressed as

$$\mathrm{TL}=20\log R+\alpha R, \tag{2}$$

where α is the absorption coefficient at a frequency of 130 kHz. For water temperatures between 15 and 21 °C, the absorption coefficient at a frequency of 130 kHz is approximately 0.038 dB/m. The source energy flux density can be expressed as

$$\mathrm{SE}=10\log\int_0^T p_s^2(t)\,dt, \tag{3}$$

where $p_s(t)$ is the acoustic pressure of the echolocation signal referenced to 1 m from the porpoise. If A is the peak amplitude of an echolocation signal, then following the procedure of Au (1993), the source acoustic pressure can be expressed as

$$p_s(t)=As(t), \tag{4}$$

where $s(t)$ is the normalized waveform and has a maximum value of 1. Inserting Eq. (4) into Eq. (3) and after some simple manipulations, the source energy flux density can be expressed as

$$\mathrm{SE}=\mathrm{SL}_{\mathrm{pp}}-6+10\log\int_0^T s^2(t)\,dt, \tag{5}$$

where $\mathrm{SL}_{\mathrm{pp}}$ is the peak-to-peak source level and is equal to $20\log 2A$. Equation (5) indicates that both the peak-to-peak source level and the energy within the normalized waveform contribute to the source energy flux density. From 1401 echolocation signals, the log (integral) term of Eq. (5) was calculated to be approximately −50.7 dB. Inserting Eqs. (2) and (5) into Eq. (1) gives the following

$$\mathrm{EE}=\mathrm{SL}_{\mathrm{pp}}-56.7-40\log R-0.07R+\mathrm{TS}_E. \tag{6}$$

Inserting the detection-threshold range of 26 m and target strength of −26 dB for the 7.62-cm sphere, and 15.9 m and −36.6 dB for the 5.08-cm sphere, gives the following:

$$\mathrm{EE}=\mathrm{SL}_{\mathrm{pp}}-\begin{cases}141.3 & \text{for 7.62-cm sphere}\\ 142.6 & \text{for 5.08-cm sphere.}\end{cases} \tag{7}$$

Equation (7) indicates that the difference in the echo energy flux density at threshold for both targets was only 1.3 dB. This shows very good agreement between the data obtained from the same animal almost one year apart.

A porpoise performing a target-detection task typically produces click trains with high degrees of variation in amplitude. This makes it difficult to determine the threshold echo energy. However, assuming that the animal is most likely to be using the clicks of high amplitude, the echo energy at threshold can be estimated by the maximum amplitude of all clicks with amplitudes within 3 dB of the maximum amplitude per trial. The porpoise was initially trained for the target-detection task using a 2.1-cm outer

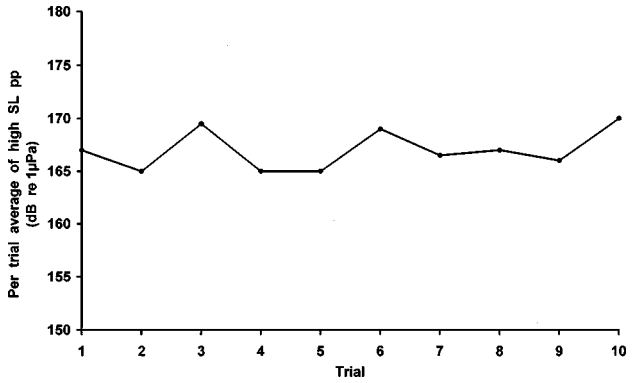


FIG. 7. Average of the peak-to-peak source level of all signals within 3 dB of the maximum amplitude per trial in ten trials.

diameter×15-cm long hollow copper tubing. During this training period, its echolocation signals and beam pattern in the horizontal and vertical planes were measured (Au *et al.*, 1998). The copper cylinder was at ranges of 7–9 m. From the maximum peak-to-peak amplitude of the emitted signals in ten trials, the average of all the signals with amplitude that were within 3 dB of the maximum was calculated (Fig. 7). Figure 7 shows that the porpoise typically used signals with peak-to-peak sound-pressure levels between 165 and 170 dB *re* 1 μPa. Therefore, the received echo energy flux density at threshold was between 22.4 and 27.4 dB *re* 1 μPa^2 s.

The detection ranges for stainless-steel spheres found in the present study could be used to estimate the distance at which an echolocating harbor porpoise can detect various gill nets, if the target strengths of such nets are measured. Au and Jones (1991) measured the target strength of a monofilament gill net constructed of 0.49-mm nylon monofilament lines with a 10-cm mesh size using simulated bottlenose dolphin (*Tursiops truncatus*) clicks. They obtained target-strength values of −52.6 dB at normal incident to −53.7 dB at a 45 degree incident angle. Au (1994) later applied the results to a harbor porpoise echolocation signal and found that the target strength would be very similar. Inserting a target strength of −52.6 dB and an energy flux density at threshold of 22.4–27.4 dB in Eq. (7) gives a detection range for the monofilament gill net of 6.5–8.5 m in a low-noise environment. However, the target strength of other types of gill nets has not been measured. To reduce bycatch, sound reflectors could be added to nets that have low detectability. The detection ranges of possible sound reflectors can also be estimated by using the equations given here.

ACKNOWLEDGMENTS

Carolien Staal and Muriël Drenthe helped train the porpoise. We thank Muriël Drenthe, Katja Wijnbergh, Sandra van Orsouw, Lydia de Boer, Helmi ter Horst, and Guusje Poot for collecting data, and Addi Kappers, Rick Wieland, and Willem Zoetebier for technical assistance. Funding for this project was obtained from: Harderwijk Marine Mammal Park (through Ruud de Clercq), The North Sea Directorate (DNZ; through Wanda Zevenboom; contract number: 76/315725) and The Netherlands National Institute for Coastal and Marine Management (RIKZ; through Erik Jagtman; contract number: 22971983), both departments of the Netherlands Ministry of Transport, Public Works, and Water Management (RWS), The Department of Nature Management (NBLF; through Jan-Willem Sneep and Gerrit van Ommering; contract numbers 9785-1997-2204-11110039 and 3003009274) of the Netherlands Ministry of Agriculture, Nature Management, and Fisheries. We thank Leen Verburg for donating the research trailer. We thank Paul Nachtigall and Jeff Pawloski of the Hawaii Institute of Marine Biology for their advice during the planning phase of the study. We thank Wim Verboom (TNO-Institute of Applied Physics, The Netherlands), Lee Miller (Odense University, Denmark), and Nancy Vaughan (University of Bristol, UK) for their constructive comments on the manuscript.

Andersen, S. (**1970**). ''Auditory Sensitivity of the Harbour Porpoise *Phocoena phocoena*,'' Invest. Cetacea **3**, 255–259.

Au, W. W. L., Moore, P. W. B., and Pawloski, D. A. (**1988**). ''Detection of Complex Echoes in Noise by an Echolocating Dolphin,'' J. Acoust. Soc. Am. **83**, 622–668.

Au, W. W. L., and Snyder, K. J. (**1980**). ''Long-range Target Detection in Open Waters by an Echolocating Atlantic Bottlenose Dolphin (*Tursiops truncatus*),'' J. Acoust. Soc. Am. **68**, 1077–1084.

Au, W. W. L., and Jones, L. (**1991**). ''Acoustic Reflectivity of Nets: Implications Concerning Incidental Take of Dolphins,'' Mar. Mamm. Sci. **7**, 258–273.

Au, W. W. L. (**1993**). *The Sonar of Dolphins* (Springer, New York).

Au, W. W. L. (**1994**). ''Sonar Detection of Gillnets by Dolphins: Theoretical Predictions,'' Rep. Inst. Whal. Commn. (Special Issue 15), pp. 565–571.

Au, W. W. L., Kastelein, R. A., Rippe, H. T., and Schooneman, N. M. (**1998**). ''The Transmission Beam Pattern and Echolocation Signal of the Harbor Porpoise (*Phocoena phocoena*),'' Proceedings of Biological Sonar Conference, Carvoeiro, Portugal 21 May–2 June 1998.

Bibikov, N. G. (**1992**). ''Auditory Brainstem Responses in the Harbor Porpoise (*Phocoena phocoena*),'' in: *Marine Mammal Sensory Systems*, edited by J. A. Thomas, R. A. Kastelein, and A. Ya. Supin (Plenum, New York), pp. 197–211.

Busnell, R. G., and Dziedzic, A. (**1967**). ''Resultants Metrologiques Experimentaux de L'echolocation chez le *Phocaena phocaena* et leur Comparaison avec Ceux de Certaines Chauves-souris,'' in *Animal Sonar System, Biology and Bionics*, edited by R. G. Busnell (Lab. Physiol. Acoust. Jouey-en-Josas, France), pp. 307–356.

Green, D. M., and Swet, J. A. (**1966**). *Signal Detection Theory and Psychophysics* (Krieger, Huntington, NY).

Jefferson, T. A., and Curry, B. E. (**1994**). ''A Global Review of Porpoise-(*Cetacea: phocoenidae*) Mortality in Gillnets,'' Biol. Conserv. **67**, 167–183.

Johnson, C. S. (**1968**). ''Relationship between Absolute Threshold and Duration-of-Tone Pulses in the Bottlenose Porpoise,'' J. Acoust. Soc. Am. **43**, 757–763.

Kamminga, C., and Wiersma, H. (**1981**). ''Investigations of Cetacean Sonar II. Acoustical Similarities and Differences in Odontocete Sonar Signals,'' Aquatic Mammals **8**, 41–62.

Møhl, B., and Andersen, S. (**1973**). ''Echolocation: High-Frequency Component in the Click of the Harbour Porpoise (*Phocoena ph. L.*),'' J. Acoust. Soc. Am. **53**, 1368–1372.

Urick, R. J. (**1983**). *Principles of Underwater Sound*, 3rd ed. (McGraw-Hill, New York).

Verboom, W. C., and Kastelein, R. A. (**1995**). ''Acoustic Signals by Harbour Porpoises (*Phocoena phocoena*), in *Harbour Porpoises, Laboratory Studies to Reduce Bycatch*, edited by P. E. Nachtigall, J. Lien, W. W. L. Au, and A. J. Read (De Spil Publishers, Woerden, The Netherlands), pp. 1–39.

Verboom, W. C., and Kastelein, R. A. (**1997**). ''Structure of Harbour Porpoise (*Phocoena phocoena*) click train signals,'' in *The Biology of the Harbour Porpoise*, edited by A. J. Read, P. R. Wiepkema, and P. E. Nachtigall (De Spil Publishers, Woerden, The Netherlands), pp. 343–362.

Zar, J. H. (**1984**). *Biostatistical Analysis*, 2nd ed. (Prentice-Hall, Englewood Cliffs, NJ).

Quantifying complex patterns of bioacoustic variation: Use of a neural network to compare killer whale (*Orcinus orca*) dialects

V. B. Deecke
Marine Mammal Research Unit, University of British Columbia, 6248 BioSciences Road, Vancouver B.C. V6T 1Z4, Canada

J. K. B. Ford
Marine Mammal Research Unit, University of British Columbia, 6248 BioSciences Road, Vancouver B.C. V6T 1Z4, Canada and Vancouver Aquarium Marine Science Centre, P.O. Box 3232, Vancouver B.C. V6B 3X8, Canada

P. Spong
OrcaLab, Hanson Island, P.O. Box 258, Alert Bay B.C. V0N 1A0, Canada

(Received 28 April 1998; revised 23 October 1998; accepted 5 January 1999)

A quantitative measure of acoustic similarity is crucial to any study comparing vocalizations of different species, social groups, or individuals. The goal of this study was to develop a method of extracting frequency contours from recordings of pulsed vocalizations and to test a nonlinear index of acoustic similarity based on the error of an artificial neural network at classifying them. Since the performance of neural networks depends on the amount of consistent variation in the training data, this technique can be used to assess such variation from samples of acoustic signals. The frequency contour extraction and the neural network index were tested on samples of one call type shared by nine social groups of killer whales. For comparison, call similarity was judged by three human subjects in pairwise classification tasks. The results showed a significant correlation between the neural network index and the similarity ratings by the subjects. Both measures of acoustic similarity were significantly correlated with the groups' association patterns, indicating that both methods of quantifying acoustic similarity are biologically meaningful. An index based on neural network analysis therefore represents an objective and repeatable means of measuring acoustic similarity, and allows comparison of results across studies, species, and time. © *1999 Acoustical Society of America.* [S0001-4966(99)01004-8]

PACS numbers: 43.80.Ka, 43.80.Lb, 43.80.Jz [FD]

INTRODUCTION

A widespread problem in the study of animal and human vocal communication lies in describing and quantifying the similarity of acoustic signals. A quantitative measure of acoustic similarity is crucial to any study comparing the vocalizations of different species, social groups, or individuals. Current approaches to this problem fall into two categories. *Statistical measures* of acoustic similarity use univariate or multivariate statistics on measures extracted from acoustic signals (e.g., Bailey, 1978; Symmes *et al.*, 1979; Clark *et al.*, 1987; Buck and Tyack, 1993; for overviews see Martindale, 1980, and Williams and Slater, 1991). *Perceptual measures* quantify acoustic similarity through ratings by human subjects (e.g., Tyack, 1986; Sayigh *et al.*, 1990), or by the ability of human or animal subjects to discriminate between classes of signals (e.g., Miller and Nicely, 1955; Loesche *et al.*, 1992).

Statistical measures of acoustic similarity have the advantage of being objective and repeatable (Martindale, 1980; Clark *et al.*, 1987), making it possible to compare the results from different studies. However, they may not always be the most meaningful, since they only assess the physical properties of the signals and give no information on how they are perceived (see Horn and Falls, 1996). Perceptual measures, although often biologically meaningful, have the problem of observer bias. Whereas ratings of similarity by the same subject are probably comparable, ratings made by different subjects are generally not. In addition, obtaining ratings by human subjects or trained animals becomes a logistic challenge in experiments where the acoustic similarity of multiple samples needs to be assessed in pairwise comparisons, or where sample sizes are large.

In this paper, we introduce the use of an artificial neural network to measure the similarity of discrete calls of killer whales (*Orcinus orca*). Artificial neural networks were developed by modeling biological systems of information processing (for overviews, see Dasgupta, 1991; Hinton, 1992). Due to their ability to classify unknown data based on information obtained from a known training set, neural networks have successfully been used in the automated classification of acoustic signals (e.g., Neumann *et al.*, 1992; Ramani *et al.*, 1993), including killer whale calls (Spong *et al.*, 1993). Since the performance of a neural network depends on the amount of consistent variation between the signal patterns in the training set, we demonstrate that the discrimination error of a neural network can be used to quantify the similarity of signals. Its bio-mimetic nature makes neural network analysis a promising candidate for a measure of similarity which assesses acoustic variation in a biological

meaningful while being objective and repeatable.

Ford (1984, 1989, 1991) showed that different killer whale communities use distinctively different vocal signals. Within the Northern Resident Community of British Columbia, stable kin groups, called pods, have unique vocal repertoires of 7–17 discrete call types. Related pods often use structurally distinct versions of the same call types. Within pods, matrilineal groups, called subpods, again have their own versions of shared call types. Finally, individuals likely have unique ''voices'' due to variation in their sound-producing structures. The vocal communication of killer whales exhibits variation on a variety of levels and provides a challenging field in which to test methods of measuring acoustic similarity.

Many studies have used frequency contours to describe vocalizations (e.g., Bailey, 1978; Sayigh *et al.*, 1990; Buck and Tyack, 1993; McCowan, 1995). For tonal signals, a frequency contour gives changes in the fundamental frequency of a vocalization over time. For pulsed signals, such as the discrete calls of killer whales (Schevill and Watkins, 1966; Ford, 1989), the contour describes changes in the pulse repetition rate (pulse-rate contour). The similarity of samples of frequency contours can be assessed using statistical (Bailey, 1978; Buck and Tyack, 1993; McCowan, 1995) or perceptual (Sayigh *et al.*, 1990) measures. Using frequency contours to describe vocalizations has the advantage that the signal is analyzed as a unit rather than broken down into disjunct measurements. In addition, irrelevant information, such as background noise or artifacts introduced by the recording apparatus, is eliminated from subsequent analyses. This is especially beneficial in the present study, which compares calls from recordings made in the field with a variety of recording systems.

So far, most automated procedures for extracting frequency contours from spectrograms have been developed for tonal signals (such as bird vocalizations or dolphin whistles) and for recordings obtained under controlled circumstances from captive or temporarily isolated animals (e.g., Buck and Tyack, 1993). In this paper we describe a method to determine the pulse repetition rate from spectrograms of pulsed calls. This method of extracting pulse-rate contours is robust to levels of background noise typical of field recordings. We introduce an index of acoustic similarity based on the performance of a neural network at classifying unknown contours using information obtained from a known training set. We test the contour extraction algorithm and the neural network index on calls of nine matrilineal groups of killer whales. For comparison, we measure the similarity of the same calls using the classification error of three human subjects. To investigate whether both measures of acoustic similarity are biologically meaningful, we compare them to the association patterns of the nine groups.

I. METHODS

A. Extraction of pulse-rate contours: The sidewinder algorithm

The discrete calls of killer whales are pulsed signals in which a tone (of a certain *tonal frequency*) is not emitted continuously but in pulses (given by the *pulse-repetition rate*; Schevill and Watkins, 1966; Watkins, 1967). Unlike in the tonal signals of many birds or other delphinids, the highest amount of energy is therefore not always contained in the first, second, or third harmonic (Watkins, 1967). The pulsed nature of these calls and the fact that the recordings used in this study were made in the field and often contained high levels of background noise meant that extraction algorithms from the literature (such as used by Buck and Tyack, 1993) proved not to be satisfactory.

For the extraction of pulse-rate contours, suitable calls were digitized at a sampling rate of 22 050 Hz from cassette tapes, including at least 100 ms of background noise before the onset of the call. Spectrograms were generated by fast Fourier transform (FFT) using the Canary 1.2.1 sound analysis software (Cornell Laboratory of Ornithology) with a filter bandwidth of 88 Hz, and an FFT size and frame length of 1024 points. Overlap between frames was 87.5%, and a Hamming window function was used for normalization. These parameters give a frequency resolution of 21.53 Hz, and a temporal resolution of 5.81 ms. Contours were extracted using MATLAB 4.2 (The MathWorks, Inc.) for Macintosh with the signal processing toolbox.

The algorithm used in this study assumes that the beginning and the end of the call can be determined visually from the spectrogram. In order to reduce background noise levels, an average noise spectrum was computed from the part of the spectrogram before the onset of the call, and subtracted from all time bins. In a spectrogram of a pulsed vocalization, the pulse repetition rate is given by the spacing between frequency bands (Watkins, 1967). To find the pulse-repetition rate at each point in time, the autocovariance sequence (mean-removed autocorrelation sequence) R was first computed for each individual power spectrum y of the spectrogram using the formula:

$$R_y(n) = \sum_{m=0}^{m=N} [y(n+m) - \overline{y}][y(n) - \overline{y}], \tag{1}$$

where n is the frequency bin number, m is the offset of the spectrum in frequency bins, $\overline{y}$ is the average sound pressure of the spectrum, and N is the number of frequency bins in the spectrum. To save computing time, the sequence was only calculated from $m=0$ to $m=N$, since the segment from $m=-N$ to $m=0$ is an exact mirror image and yields no additional information. The frequencies of any sidebands in the acoustic signal are given by a simple linear relationship, and therefore the autocovariance sequence will show a peak every time m equals a multiple of the frequency spacing of the bands (i.e., of the pulse-repetition rate; Watkins, 1967) and adjacent bands overlap. Because the power spectrum of the background noise tends to decrease with increasing frequency, and adjacent frequency bands generally have similar energy content, the second highest maximum in the autocovariance sequence (after $m=0$) usually corresponds to the frequency bin containing the pulse-repetition rate. Sometimes this maximum represents the second, and in some rare cases the third, harmonic. A simple heuristic algorithm described by Buck and Tyack (1993), which checked for local

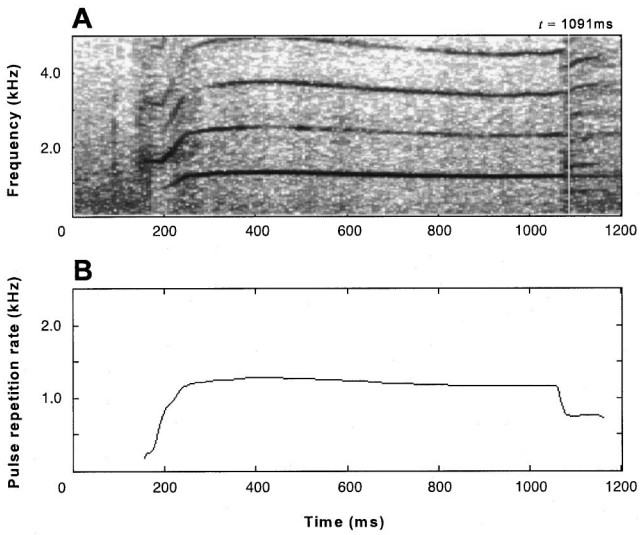


FIG. 1. (A) Spectrogram of an N4 call with a filter bandwidth of 88 Hz. The white line at $t = 1091$ ms shows the position of the power spectrum in Fig. 2. (B) Pulse-rate contour extracted from the spectrogram.

maxima at 1/2 and 1/3 of the offset of the second highest maximum, could account for this.

Figure 1 shows a spectrogram of an N4 call and a pulse-rate contour extracted from it. Figure 2 gives the power spectrum at $t = 1091$ ms (A) and its covariance sequence (B) for the same call. For subsequent analysis, the pulse-repetition rate was determined at 100 equally spaced points throughout the call and presented to the neural network as a vector of 100 numbers. Thus calls were essentially standardized for time; however, call length was entered as a 101 st number

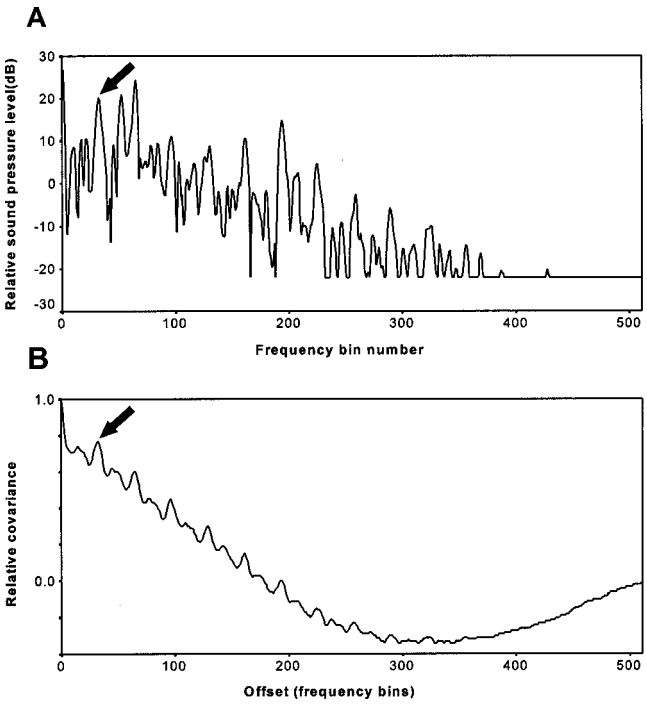


FIG. 2. (A) Power spectrum of the terminal component of an N4 call at $t = 1091$ ms (see Fig. 1). Filter bandwidth is 88 Hz and frequency resolution is 21.53 Hz per frequency bin. (B) Autocovariance sequence of the power spectrum. The arrows indicate the frequency bin containing the pulse-repetition rate.

into the analysis to allow discrimination of calls which differed consistently in length, but not in structure.

B. Analysis of acoustic variation in the N4 call

To test the performance of the neural network index on biological data, we used recordings of nine matrilineal groups, or subpods, of killer whales. Matrilineal groups consisted of between two and seven individuals, and belong to A-subclan of the Northern Resident Community (Ford, 1991) which inhabits the waters of British Columbia, Canada. The recordings analyzed in this study were made in the fjords and straits of the southern coast of British Columbia in weakly stratified or unstratified waters of depths of up to 400 m and often contained low to moderate levels of shipping noise. Recordings were contributed by a number of researchers using a variety of recording systems. All systems had a flat frequency response from 0.1 to 7 kHz, although for some systems the range of the flat response extended up to 20 kHz.

All members of the Northern Resident Community can be identified consistently from natural markings (Bigg *et al.*, 1990; Ford *et al.*, 1994). The analysis was restricted to recordings which could be attributed to a certain matrilineal group because was the only group within recording range and its identity was confirmed visually or photographically. We chose the N4 call (Ford, 1989, 1991; see Fig. 3) for this study because it is shared by all nine groups and it is one of the most frequently used call types in their repertoire. Structurally, the N4 calls of A08, A09, A23, A25, and A36 all have relatively low peak pulse repetition rates, and a pronounced terminal component at the end of the call (see Fig. 3). The versions of N4 made by A12 and A30 subpods usually lack the terminal component and have a relatively higher peak pulse repetition rate. Finally the N4 calls of A11 and A24 subpods (A4 pod of Ford, 1991) tend to be longer than those of any other matrilineal group and generally end in an upsweep.

N4 calls with adequate signal-to-noise ratios were identified acoustically and visually from recordings, and were digitized using the Canary 1.2.1 sound analysis software. Spectrograms were computed and pulse-rate contours extracted with the sidewinder algorithm. Since the performance of a neural network is highly dependent on the number of examples for each signal pattern in the training set, sample size for all matrilineal groups was standardized to 24, the size of the smallest sample. For each group we included calls from as many independent recording sessions as possible, to present the neural network and the human subjects with calls from a wide range of behavioral contexts, which are known to affect call structure (Ford, 1989). No less than three independent recording sessions were used for any one matrilineal group.

Association patterns of the different matrilineal groups were analyzed by generating an association matrix giving the *half-weight index of association* (Ginsberg and Young, 1992) for each pair of matrilineal groups. This index gives the number of observations of two groups traveling together as a proportion of half the total number of observations for the two groups. The association data came from a sightings da-

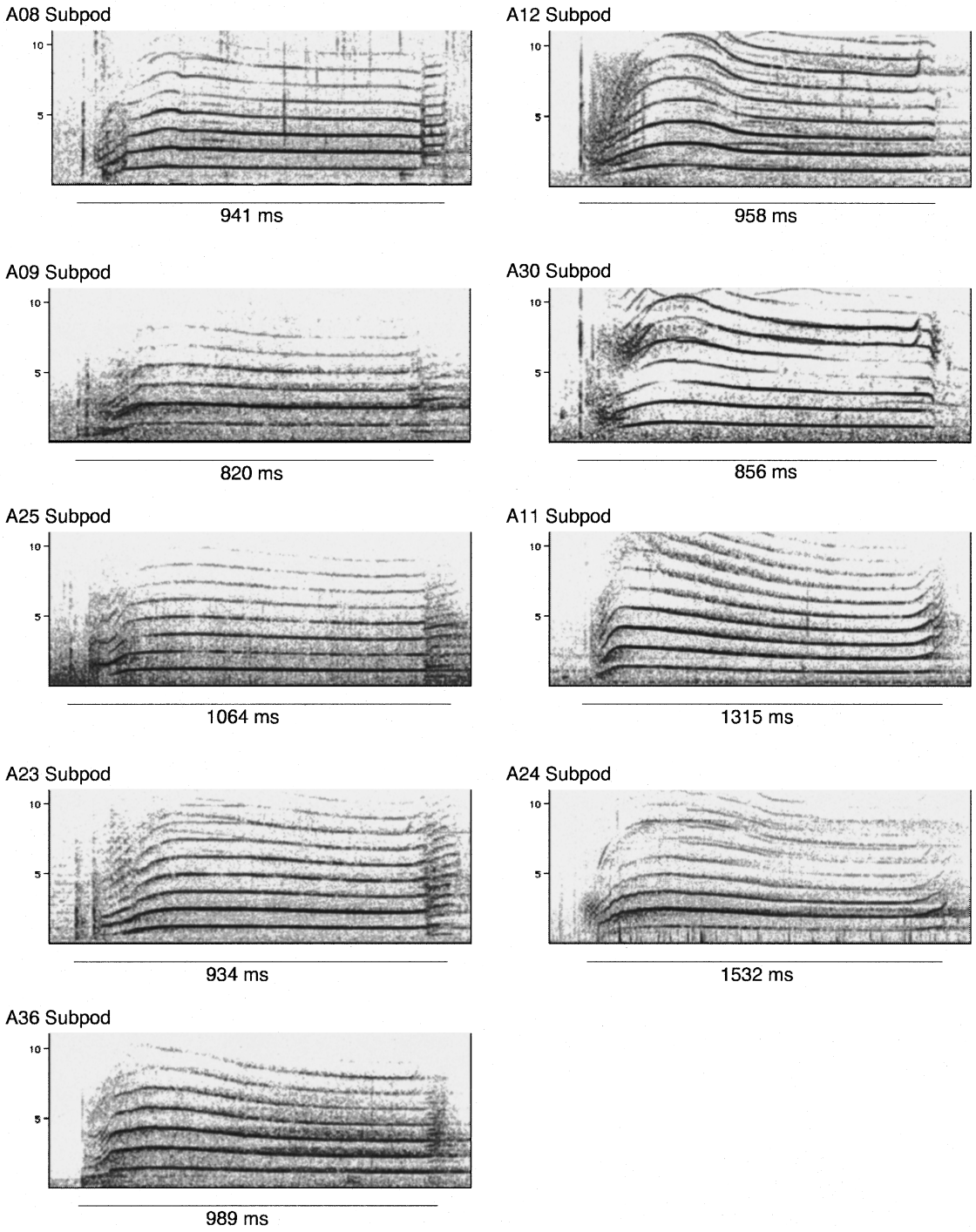


FIG. 3. Examples of spectrograms of N4 calls from the nine matrilineal groups of A-subclan.

tabase for the Northern Resident Community spanning the years 1990–1995. The total number of sightings of one or more A-subclan matrilineal group was 757, while numbers of sightings of any one matrilineal group ranged between 147 (A25 subpod) and 415 (A30 subpod).

C. The neural network index of acoustic similarity

Neural network analysis was done with the neural network toolbox of MATLAB 4.2 for the Macintosh (The MathWorks, Inc.). We used a standard back-propagation network (e.g., Rumelhart *et al.*, 1986) with momentum and an adaptive learning rate (Vogl *et al.*, 1988). Back-propagation networks can be trained to classify unknown patterns by ''learning'' to associate certain known input patterns with certain outputs. In our case inputs consisted of pulse-rate contours plus call length from two social groups, and the expected outputs were the matrices [0 1] and [1 0], depending on which group the contour came from. After training, the performance of a neural network can be tested by presenting it with data not used in training, and determining how closely the observed output matches the expected one.

To determine the network's performance during the training process, the training algorithm computes the sum-

square error of observed against expected output. We used a modified version of this parameter, the *discrimination error*, to determine network performance when testing a network with unknown contours. The discrimination error is calculated by subtracting the observed output from the expected output ([0 1] or [1 0]) and taking the mean of the absolute differences. The average of the discrimination error of all networks trained on one comparison was chosen over the proportion of misclassifications because the discrimination error not only yields information on whether a classification is correct, but also gives a relative measure of the quality of discrimination. For example, even an untrained network might easily generate [0.49 0.51] for an expected output of [0 1]. Since in both cases the second output exceeds the first one, the classification is considered correct and the neural network has a classification error of 0, although the classification was hardly better than random. The discrimination error of 0.49 much better reflects the poor quality of this classification.

The optimal neural network architecture for the discrimination tasks was determined in a network design experiment which consisted of training neural networks on a range of comparisons and varying the number of neurons in the hidden layer, as well as the length of training. For all comparisons, discrimination did not increase detectably when using more than 20 neurons in the hidden layer and when training for more than 5000 iterations, so that these parameters were used in all subsequent analyses. Neural networks were initialized with random weights, and a small number of neural networks did not improve in performance from the initialized state. Since this failure to train results from the configuration of weights at initialization rather than from a lack of consistent variation in the training set, such networks were eliminated from the analysis by setting a criterion of a 20% decrease of the sum-square error during the first 150 iterations.

To arrive at an acoustic similarity matrix for the N4 calls of the nine groups, we trained and tested neural networks on all 36 possible pairwise comparisons. We intended to train as many independent neural networks as possible on each comparison to eliminate the stochastic component of neural network analysis. To do this, one contour was excluded from the training set, a neural network was trained on the remaining 47 contours, the neural network was tested using the excluded contour and the discrimination error was determined. The test contour was then added back to the training set, another one was removed, and this procedure was repeated until each contour had served as the test contour. We tested each network with only a single contour in order to have as many contours for training as possible. Networks trained with fewer contours and tested with more gave consistently higher discrimination errors, probably due to overtraining (Rumelhart *et al.*, 1986). The *neural network index of acoustic similarity* for each pairwise comparison is the average of the discrimination errors of all 48 neural networks trained this way.

D. Acoustic similarity ratings by human subjects

The same nine samples of 24 calls each were used to determine the classification errors of human subjects in pairwise computer-based discrimination tasks. Three female subjects, none of whom had any previous knowledge of killer whale vocal communication, were presented with the discrimination tasks. Subject A was 20 years old and had no musical background. Subject B was 22 years old and had played the flute for 2 years, and subject C was 20 years old and had played the piano for 13 years.

Since human subjects cannot be trained more than once on the same problem without seeing an increase in performance, we used a somewhat modified training and testing protocol for this part of the analysis. In each discrimination task, the subject was first presented with a training set of 16 calls belonging to two categories (A or B) according to the group they came from. The subject could listen to the calls and view their spectrograms, and was then asked to assign a test set of 32 unfamiliar calls to the appropriate category. The *rating of acoustic similarity* gives the proportion of misclassifications among these 32 calls. During the testing, the subject was allowed to return to the training set, but in order to cause her to generalize, was asked not to do so more than three times for each discrimination task. Following the experiments, the subjects completed a questionnaire asking whether they classified the calls primarily using acoustic or visual cues.

For visual comparison, average linkage dendrograms were generated from the four acoustic similarity matrices (one neural network index and three human subject ratings) as well as from the association matrix. Average linkage is a hierarchical tree-building algorithm and will group subpods with high indices of acoustic similarity or association into common clusters in a dendrogram (see Johnson, 1967). The acoustic similarity matrices and the association matrix were compared statistically by generating the matrix correlation coefficient for all possible comparisons. A Mantel test was used to test for significance.

II. RESULTS

The sidewinder algorithm proved effective at extracting pulse-rate contours from recordings obtained under a variety of recording conditions. Contours could be obtained even from recordings with high levels of ambient noise, if the call was clear and the energy in two or more frequency bands exceeded the background noise level. Only recordings containing boat noise with harmonic content, and recordings with a great amount of acoustic reverberation or strong echoes, caused problems in the contour extraction.

The values for the neural network index of acoustic similarity for the pairwise comparisons of N4 calls are given in Table I. The neural network could best discriminate between the N4 calls of A23 and A24 subpods (neural network index: 0.01). A09 subpod and A25 subpods gave the poorest discrimination (neural network index: 0.48). The average value for the neural network index for all discrimination tasks was 0.15. The neural network index grouped the nine matrilineal groups into three major clusters according to the

TABLE I. Acoustic similarity matrix for the N4 call of the nine matrilineal groups based on the neural network index of acoustic similarity. The values give the neural network performance (average discrimination error) for each pairwise comparison.

A09	0.34							
A11	0.09	0.04						
A12	0.05	0.03	0.08					
A23	0.18	0.19	0.04	0.04				
A24	0.10	0.03	0.29	0.08	0.01			
A25	0.43	0.48	0.04	0.05	0.23	0.03		
A30	0.14	0.06	0.07	0.37	0.09	0.10	0.06	
A36	0.27	0.40	0.07	0.06	0.19	0.06	0.37	0.10
	A08	A09	A11	A12	A23	A24	A25	A30

similarity of their N4 calls. These are A08–A09–A23–A25–A36, A12–A30, and A11–A24 (Fig. 4). These clusters are consistent with structural differences in the calls shown in Fig. 3.

Table II gives the ratings of acoustic similarity (proportion of misclassifications) for the 3 subjects and the 36 classification tasks. The table shows that subjects B and C classified all calls correctly in at least one comparison. The highest proportion of misclassification was higher than random (0.63, A08 vs A09 by subject C). The average proportion of misclassification for all discrimination tasks was 0.25, 0.18, and 0.15 for subjects A, B, and C, respectively, and a

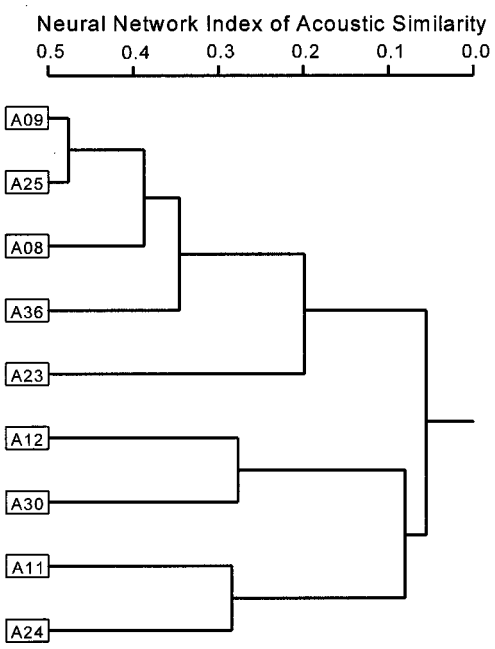


FIG. 4. Average linkage dendrogram giving acoustic similarity of the N4 call of the nine matrilineal groups based on the neural network index (generated from Table I). The position of the vertical lines linking groups or clusters of groups with respect to the scale bar above indicates the similarity of their N4 call based on the neural network index. Comparisons of N4 calls from groups which are linked on the left-hand side of the graph gave higher average discrimination errors (suggesting higher similarity) than those linked on the right-hand side.

TABLE II. Acoustic similarity matrix for the N4 call of the nine matrilineal groups generated by three human subjects. The values give the subjects' performance (proportion of misclassifications) for each pairwise comparison.

	Subject								
	A	0.44							
A09	B	0.25							
	C	0.63							
	A	0.13	0.22						
A11	B	0.25	0.06						
	C	0.16	0.03						
	A	0.16	0.06	0.06					
A12	B	0.09	0.09	0.09					
	C	0.06	0.09	0.03					
	A	0.50	0.25	0.03	0.25				
A23	B	0.19	0.47	0.03	0.09				
	C	0.19	0.53	0.00	0.03				
	A	0.25	0.28	0.44	0.28	0.03			
A24	B	0.16	0.06	0.38	0.22	0.03			
	C	0.19	0.06	0.28	0.00	0.03			
	A	0.44	0.50	0.03	0.22	0.41	0.16		
A25	B	0.44	0.50	0.13	0.09	0.38	0.03		
	C	0.25	0.34	0.13	0.00	0.38	0.00		
	A	0.13	0.13	0.22	0.31	0.09	0.38	0.06	
A30	B	0.09	0.13	0.09	0.38	0.00	0.03	0.06	
	C	0.06	0.13	0.00	0.41	0.00	0.00	0.00	
	A	0.34	0.25	0.22	0.34	0.44	0.22	0.50	0.38
A36	B	0.09	0.28	0.16	0.16	0.16	0.13	0.31	0.31
	C	0.22	0.19	0.13	0.13	0.28	0.06	0.25	0.28
		A08	A09	A11	A12	A23	A24	A25	A30

sign test showed that subject A made significantly more misclassification than the other two subjects ($p<0.001$). The proportions of misclassification of the three subjects for any one comparison differed on average by 0.11, and these differences ranged from 0 to 0.31. Subjects A and C said that they used mainly acoustic and some visual cues to do the discrimination, subject B said she relied mainly on the spectrogram, with some acoustic cues. Figure 5 shows that all three subjects grouped the calls of the nine matrilineal groups into three major clusters which correspond to the clusters generated by the neural network index (Fig. 4). However, the results from individual subjects differ in the relationship of matrilineal groups within the three clusters, as well as in the positions of the clusters with respect to each other.

The association matrix for the nine matrilineal groups is given in Table III. Association indices range from 0.14 for A09 and A11 subpod to 0.95 for A11 and A24 subpod. The average linkage dendrogram (Fig. 6) shows that their association patterns group the nine matrilineal groups into the same three clusters as the acoustic analyses, with the difference that the A36 subpod associates more often with A12 while being acoustically more similar to A08–A09–A23–A25.

Table IV gives the correlation matrix of the ratings of

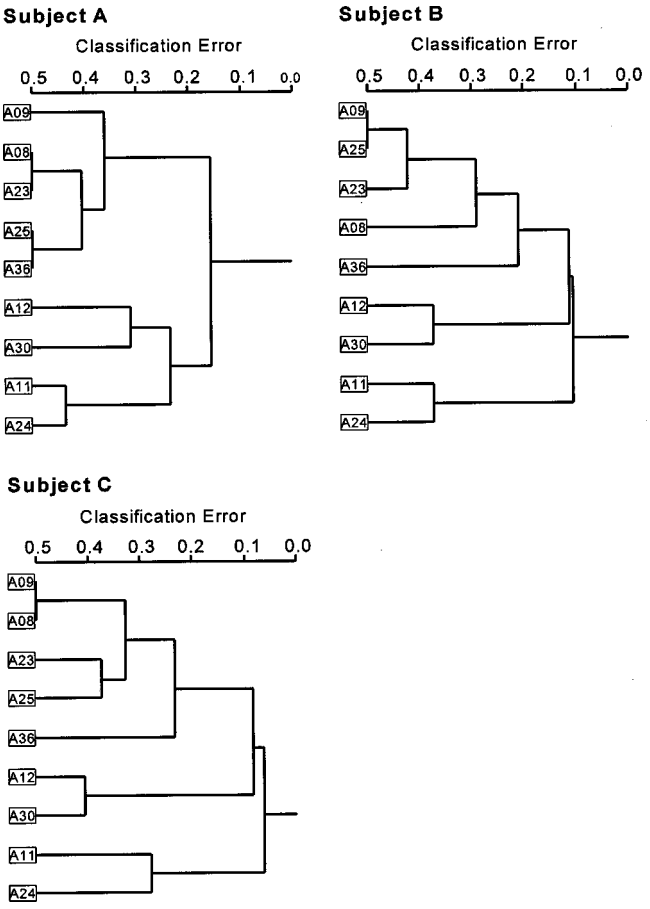


FIG. 5. Average linkage dendrogram giving acoustic similarity of the N4 call of the nine matrilineal groups based on the classification errors of the three subjects (generated from Table II). Comparisons of N4 calls from groups which are linked on the left-hand side of the graph gave higher classification errors (suggesting higher similarity) than those linked on the right-hand side

acoustic similarity by the three subjects, by the neural network, as well as of the association indices. All correlations are significant with $p<0.001$ (Mantel's test). Two correlation coefficients comparing ratings of different subjects (subjects A and B; subjects A and C) are lower than the correlation coefficients comparing human subject ratings and the neural network indices. All measures of acoustic similarity gave significant correlations with the groups' association indices.

TABLE III. Association matrix for the nine matrilineal groups. The values give the half-weight index of association.

A09	0.73							
A11	0.42	0.24						
A12	0.51	0.40	0.41					
A23	0.77	0.69	0.33	0.47				
A24	0.42	0.25	0.97	0.41	0.32			
A25	0.69	0.61	0.27	0.37	0.78	0.26		
A30	0.33	0.25	0.33	0.54	0.35	0.33	0.27	
A36	0.40	0.38	0.34	0.52	0.39	0.33	0.27	0.37
	A08	A09	A11	A12	A23	A24	A25	A30

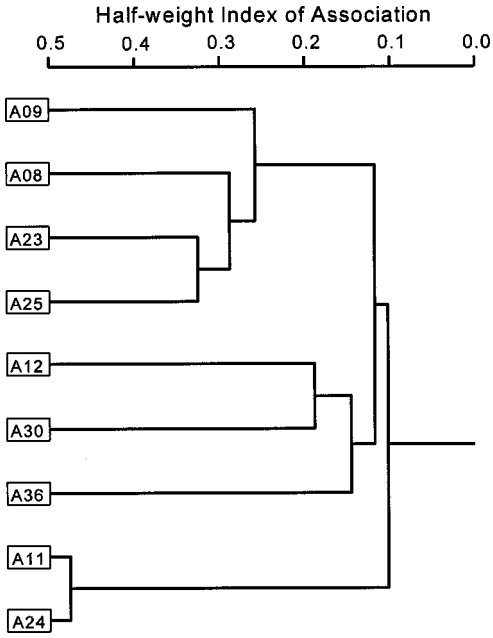


FIG. 6. Average linkage dendrogram giving association patterns of the nine matrilineal groups based on the half-weight index of association (generated from Table III). Groups which are linked on the left-hand side of the graphs spend more time traveling together than those linked on the right-hand side.

III. DISCUSSION

The contour extraction algorithm based on autocovariance in the frequency domain proved good at extracting pulse-rate contours even from recordings with poor signal-to-noise ratios. Unless the noise itself had harmonic content, it was canceled out in the autocovariance sequence, whereas the harmonic signals were amplified. We suggest that pulse-rate contours are an effective way to describe pulsed vocalizations and believe that this algorithm would be useful for extracting contours from noisy recordings of the pulsed calls of a wide variety of species.

The shortcomings of this algorithm are that it cannot be applied to broadband or pure-tone signals, and that compared to alternative algorithms, it is computationally expensive. Mixed signals, however, can still be analyzed by switching to another algorithm (e.g., that of Buck and Tyack, 1993) if the autocovariance sequence fails to detect harmonic content. Recent developments in computer hardware are likely to further reduce computing time, making real-time extraction of pulse-rate contours a possibility.

The advantage of analyses of acoustic similarity based on frequency contours over those based on isolated measurements of the spectrogram lies in the fact that analysis of frequency contours requires no, or very little, prior knowledge of where to expect the differences in the signals (Bailey, 1978). Subtle and very localized differences between two signal patterns are easily missed in conventional analyses by taking measurements of a limited number of structural variables. Unlike discrimination and classification analyses of bioacoustic signals where the input is the waveform (e.g., Neumann *et al.*, in press), or the spectrogram (e.g.

TABLE IV. Correlation matrix giving matrix correlation coefficients for the ratings of acoustic similarity by the human subjects, the neural network index, and the association indices of the nine matrilineal groups. All correlations are significant with $p<0.001$ (Mantel test).

Similarity ratings					
Subject A	1				
Subject B	0.63	1			
Subject C	0.60	0.79	1		
Neural network index	0.69	0.78	0.71	1	
Index of association	0.57	0.67	0.66	0.54	1
	Subject A	Subject B	Subject C	Neural network index	Index of association
		Similarity ratings			

Spong *et al.*, 1993; Erbe *et al.*, in press), amplitude information is excluded from the analysis of frequency contours. Although this may be a disadvantage in some studies analyzing recordings obtained in controlled environments, it will prove beneficial in others where differences in recording equipment and in the composition of background noise introduces spurious variability into the data. In the study of Spong *et al.* (1993), for example, it cannot be ruled out that the neural network discriminated along differences in background noise composition rather than individual-specific vocal differences.

Although the ratings of similarity by human subjects agreed on a fundamental level, this study suggests that individual human subjects perceive similarity of killer whale calls differently. The ratings of similarity disagree between subjects in the acoustic relationships of matrilineal groups within the three clusters, as well as in the position of these clusters with respect to each other. The subject who had never played a musical instrument had significantly higher classification errors than the other two subjects, which may suggest that the amount of musical exposure contributes to observer bias (see also Halpern *et al.*, 1995; Baribeau *et al.*, 1996; Halpern *et al.*, 1996).

Comparing the ratings of acoustic similarity by the neural network with those of the human subjects shows that both ways of quantifying acoustic variation gave similar results. The matrix correlation coefficients (Table IV) suggest that the differences between ratings from individual subjects are greater than are differences between subject ratings and the neural network index. Since multiple independent neural networks are trained on the same problem in each comparison, the neural network index will give essentially identical results given the same input data. The neural network therefore represents an objective and repeatable means of measuring acoustic similarity, and allows the comparison of results across studies, species, and time.

Like discriminant function analysis (e.g., Job *et al.*, 1995), or analysis of confusion frequencies (e.g., Miller and Nicely, 1955; Loesche *et al.*, 1992), the neural network index of acoustic similarity is based on the premise that similarity and discrimination are inversely related. All three methods rate patterns as similar if the analysis is unable to tell them apart, and conversely consider patterns distinct if the analysis can consistently discriminate between them. This concept of similarity differs from that underlying other methods which use the geometric distance between two patterns as a measure of their similarity. Examples for the latter are cross correlation (e.g., Clark *et al.*, 1987), and cluster analysis (e.g., McCowan, 1995). Arguably the first concept of similarity is more applicable to the study of communication, since the information value of a signal is largely determined by how well a receiving animal can distinguish it from other signals (Beecher, 1989).

The training procedure, which involves error back-propagation to discriminate between different patterns, is essentially a self-organizing process and does not depend on strictly linear relationships in the input data. For this reason a neural network index will be able to detect and integrate differences between the input patterns that would be missed by most conventional statistical analyses. Research into biological neural systems suggests that these also operate in a nonlinear and self-organizing way (Kelso, 1995), which may explain why a neural network based approach is often the best way to model biological signal processing tasks (Hunt, 1993; Erbe *et al.*, in press). The fact that the neural network index of acoustic similarity shows a significant correlation with the association patterns of the different matrilineal groups suggests that the index rates acoustic similarity in a biologically meaningful way.

An index of acoustic variation based on neural network analysis can be viewed as a hybrid between statistical and perceptive approaches of measuring acoustic similarity. It combines the objectivity and repeatability of a strictly statistical approach with the self-organizing nonlinear nature of acoustic perception and biological signal processing, and therefore holds great potential in the study of human and animal communication.

IV. CONCLUSIONS

This study demonstrates that autocovariance in the frequency domain is a useful way to extract contours of the pulse-repetition rate from noisy recordings of pulsed signals. This study also shows that discrimination of frequency contours using a back-propagation neural network is an effective and repeatable way to measure the similarity of animal sounds. The significant correlation between the neural network based acoustic similarity index and a biological param-

eter, the groups' association patterns, suggests that the index assesses acoustic similarity in a biologically meaningful way.

ACKNOWLEDGMENTS

This study would not have been possible without the contribution of recordings by D.E. Bain, J. Borrowman, D. Briggs, G. Ellis, B. and D. Mackay, A. Morton, F. Thomsen, and S. Wischniowski, and we are very grateful for their generosity. We would like to thank L. Farroway, S. Midwinter, and C.A. Shankel for acting as research subjects for the discrimination tasks, as well as D.E. Bain, J.M. Gosline, J.M.N. Smith, and H. Symonds for their suggestions and comments throughout the study. Earlier drafts of this article benefited greatly from comments by L. Barrett-Lennard, L. Deecke, V. M. Janik, D. Kapan, P.J.B. Slater, H. Symonds, and R.M. Williams, and two anonymous reviewers. We would like to thank The MathWorks, Inc. for generously donating the software and DKMC for logistic support.

Bailey, K. (**1978**). ''The structure and variation of the separation call of the bobwhite quail (*Colinus virginianus*, Odontophorinae),'' Anim. Behav. **26**, 296–303.

Baribeau, J., Berman, B., Atkin, A., and Roth, R. M. (**1996**). ''Musical experience and auditory P300 in a divided attention task,'' Brain Cogn. **30**, 378–380.

Beecher, M. D. (**1989**). ''Signalling systems for individual recognition: an information theory approach,'' Anim. Behav. **38**, 248–261.

Bigg, M. A., Olesiuk, P. F., and Ellis, G. M. (**1990**). ''Social organization and genealogy of resident killer whales (*Orcinus orca*) in the coastal waters of British Columbia and Washington State,'' Rep. Int. Whaling Comm. Spec. Issue No. **12**, 383–405.

Buck, J. R., and Tyack, P. L. (**1993**). ''A quantitative measure of similarity for *Tursiops truncatus* signature whistles,'' J. Acoust. Soc. Am. **94**, 2497–2506.

Clark, C. W., Marler, P., and Beeman, B. (**1987**). ''Qualitative analysis of animal vocal phonology and application to swamp sparrow song,'' Ethology **76**, 101–115.

Dasgupta, C. (**1991**). ''Neural networks: An overview,'' J. Indian Inst. Sci. **71**, 491–502.

Deng, L. (**1992**). ''Processing of acoustic signals in a cochlear model incorporating laterally coupled suppressive elements,'' Neural Networks **5**, 19–34.

Erbe, C., King, A. R., Yedlin, M., and Farmer, D. M. (**in press**). ''Computer models for masked hearing experiments with beluga whales (*Delphinapterus leucas*),'' J. Acoust. Soc. Am.

Ford, J. K. B. (**1984**). ''Call traditions and vocal dialects of killer whales (*Orcinus orca*) in British Columbia,'' Ph.D. Dissertation (University of British Columbia, Vancouver).

Ford, J. K. B. (**1989**). ''Acoustic behaviour of resident killer whales (*Orcinus orca*) off Vancouver Island, British Columbia,'' Can. J. Zool. **67**, 727–745.

Ford, J. K. B. (**1991**). ''Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters of British Columbia,'' Can. J. Zool. **69**, 1454–1483.

Ford, J. K. B., Ellis, G. M., and Balcomb, K. C. (**1994**). *Killer Whales—the Natural History and Genealogy of Orcinus orca in British Columbia and Washington State* (UBC Press, Vancouver).

Ginsberg, J. R., and Young, T. P. (**1992**). ''Measuring associations between individuals or groups in behavioural studies,'' Anim. Behav. **44**, 377–379.

Halpern, A. R., Bartlett, J. C., and Dowling, W. J. (**1995**). ''Aging and experience in the recognition of musical transpositions,'' Psychol. Aging **10**, 325–342.

Halpern, A. R., Kwak, S., Bartlett, J. C., and Dowling, W. J. (**1996**). ''Effects of aging and musical experience on the representation of tonal hierarchies,'' Psychol. Aging. **11**, 235–246.

Hinton, G. E. (**1992**). ''How neural networks learn from experience,'' Sci. Am. **268**, 145–151.

Hunt, E. (**1993**). ''A proposal for computer modelling of animal linguistic comprehension,'' in *Language and Communication: Comparative Perspectives*, edited by H. L. Roitblat, L. M. Herman, and P. E. Nachtigall (Lawrence Erlbaum, Hillsdale, NJ), pp. 85–94.

Horn, A. G., and Falls, J. B. (**1996**). ''Categorization and the design of signals: the case of song repertoires,'' in *Ecology and Evolution of Acoustic Communication in Birds*, edited by D. E. Kroodsma and E. H. Miller (Comstock Publishing, Ithaca, NY), pp. 121–135.

Job, D. A., Boness, D. J., and Francis, J. M. (**1995**). ''Individual variation in nursing vocalizations of Hawaiian monk seal pups, *Monachus schauinslandi* (Phocidae, Pinnipedia), and lack of maternal recognition,'' Can. J. Zool. **73**, 975–983.

Johnson, S. C. (**1967**). ''Hierarchical clustering schemes,'' Psychometrika **32**, 241–54.

Kelso, J. A. S. (**1995**). *Dynamic Patterns: The Self-Organization of Brain and Behavior* (MIT Press, Cambridge, MA).

Loesche, P., Beecher, M. D., and Stoddard, P. K. (**1992**). ''Perception of cliff swallow calls by birds *Hirundo pyrrhonota* and *Sturnus vulgaris* and humans *Homo sapiens*,'' J. Comp. Psych. **106**, 239–247.

McCowan, B. (**1995**). ''A new quantitative technique for categorizing whistles using simulated signals and whistles from captive bottlenose dolphins (Delphinidae, *Tursiops truncatus*),'' Ethology **100**, 177–193.

Martindale, S. (**1980**). ''On the multivariate analysis of avian vocalizations,'' J. Theor. Biol. **83**, 107–110.

Miller, E. H. (**1979**). ''An approach to the analysis of graded calls of birds,'' Behav. Neural Biol. **27**, 25–38.

Miller, G. A., and Nicely, P. E. (**1955**). ''An analysis of perceptual confusions among some English consonants,'' J. Acoust. Soc. Am. **27**, 338–352.

Neumann, E. K., Wheeler, D. A., Bernstein, A. S., Burnside, J. W., and Hall, J. C. (**1992**). ''Artificial neural network classification of *Drosophila* song mutants,'' Biol. Cybern. **66**, 485–469.

Ramani, N., Hanson, W. G., Patrick, P. H., and Sheehan, R. W. (**1993**). ''An automated environmental monitoring system—phase 1. amphibian species identification from calls,'' Proceedings of the World Conference on Neural Networks **1**, 304–307.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (**1986**). ''Learning internal representation by error propagation,'' in *Parallel Data Processing, Vol. 1*, edited by D. E. Rumelhart and J. McLelland (MIT Press, Cambridge, MA), pp. 318–362.

Sayigh, L. S., Tyack, P. L., Wells, R. S., and Scott, M. D. (**1990**). ''Signature whistles of free-ranging bottlenose dolphins *Tursiops truncatus*: Stability and mother-offspring comparisons,'' Behav. Ecol. Sociobiol. **26**, 247–260.

Schevill, W. E., and Watkins, W. A. (**1966**). ''Sound structure and directionality in *Orcinus* (killer whale),'' Zoologica **51**, 70–76.

Spong, P., Symonds, H., Gaetz, W., Jantzen, K., and Weinberg, H. (**1993**). ''A neural network method for recognition of individual *Orcinus orca* based on their acoustic behavior: Phase 1,'' Paper presented at Oceans 1993, Conference of the IEEE, Victoria, B.C.

Symmes, D., Newman, J. D., Talmage-Riggs, G., and Katz Lieblich, A. (**1979**). ''Individuality and stability of isolation peeps in squirrel monkeys,'' Anim. Behav. **27**, 1142–1152.

Tyack, P. L. (**1986**). ''Whistle repertoires of two bottlenose dolphins (*Tursiops truncatus*): Mimicry of signature whistles?'' Behav. Ecol. Sociobiol. **18**, 251–257.

Vogl, T. P., Mangis, J. K., Ziegler, A. K., and Alkon, D. L. (**1988**). ''Accelerating the convergence of the backpropagation method,'' Biol. Cybern. **59**, 257–263.

Watkins, W. A. (**1967**). ''The harmonic interval: fact or artifact in spectral analysis of pulse trains,'' in *Marine Bioacoustics, Vol. 2*, edited by W. N. Tavolga (Pergamon New York), pp. 15–43.

Williams, J. M., and Slater, P. J. B. (**1991**). ''Computer analysis of bird sounds: a guide to current methods,'' Bioacoustics **3**, 121–128.

Internal deformation of a uniform elastic solid by acoustic radiation force

William F. Walker
Department of Biomedical Engineering, The University of Virginia, Stacey Hall, 1105 West Main Street, Charlottesville, Virginia 22903

(Received 18 March 1998; revised 9 July 1998; accepted 12 January 1999)

Tissue elasticity estimation is a growing area of ultrasound research. One proposed approach would apply acoustic radiation force to displace tissue and use ultrasonic motion tracking techniques to measure the resultant displacement. Such a technique might allow noninvasive imaging of tissue elastic properties. The potential of this method will be limited by the magnitude of displacements which can be generated at reasonable acoustic intensity levels. This paper presents methods for estimating the internal displacements induced in an elastic solid by acoustic radiation force. These methods predict displacements on the order of 400 μm in the human vitreous body, 0.008 μm in human breast, and 0.020 μm in human liver at an acoustic intensity of 1.0 W/cm^2 (in water) and an operating frequency of 10 MHz. While the displacement generated in the vitreous should be readily detectable using ultrasonic methods, the displacements generated in the breast and liver will be much more difficult to detect. Methods are also developed for predicting the time dependent temperature increases associated with attenuated acoustic fields in the absence of perfusion. These results indicate promise for radiation force imaging in the vitreous, but potential difficulties in applying these techniques in other parts of the body. © *1999 Acoustical Society of America.* [S0001-4966(99)03604-8]

PACS numbers: 43.80.Qf [FD]

INTRODUCTION

Tissue elasticity estimation is a rapidly growing area of ultrasound research. Most proposed methods apply either constant[1,2] or sinusoidally varying[3–5] external displacement fields and track induced displacements using a variety of ultrasonic methods. A similar technique has been proposed which applies an external vibration, but utilizes Magnetic Resonance Imaging (MRI) to track resulting displacements.[6] While these methods have shown potential, none have gained widespread use, possibly because the complex interaction between material properties and boundary conditions makes it difficult to measure material properties, or because of the presence of a variety of artifacts. Furthermore, since these methods couple external forces into the tissue, they may not be well suited to imaging soft tissues within harder shells, such as the brain and eye.

Recently acoustic radiation force has been proposed as a method of interrogating tissue elastic properties.[7–13] In one implementation, two ultrasonic beams with slightly different center frequencies are overlaid and a beat frequency is generated. The amplitude of this beat frequency is measured using an external microphone.[8] A map of tissue response is made as the ultrasonic beams are scanned throughout the tissue. In an alternate approach, a high intensity ultrasonic beam is modulated so that shear waves propagate out from the beam axis.[11,12] The displacements induced by these shear waves are then estimated.

In another approach, a series of ultrasonic pulses are transmitted in the same direction and the resultant tissue displacement is measured as a function of pulse number.[9,10,13,14] Since this approach measures the tissue displacement as a function of time, it can be used to detect the presence of fluid,[13,14] estimate tissue elasticity,[9,10,13] or estimate tissue viscoelastic properties.[9,10] Multiple beam lines can be interrogated sequentially to form two- or even three-dimensional images. The ultimate resolution and sensitivity of this method will depend upon the magnitude and shape of the displacement field induced by the applied acoustic radiation force. In some applications a highly localized force field may be desirable to generate localized displacements to avoid artifacts in the displacement field caused by fixed tissue boundaries. In other applications a broad force field may be preferable to allow rapid image generation with possibly greater displacements.

This paper uses analytical and numerical methods to predict the magnitude and shape of tissue displacement fields induced by the application of acoustic radiation force. These methods are applied to predict the displacements which might be induced by radiation force in the human vitreous, liver, and breast. Methods are also developed to predict the time dependent heating associated with the attenuation of acoustic fields in the absence of perfusion. Finally, conclusions regarding the potential of radiation force imaging are presented.

DISPLACEMENT ANALYSIS

We begin the analysis of the displacement field induced by acoustic radiation force by considering the displacement field induced in an infinite isotropic medium by a point force acting in the positive z direction at the origin:[15]

$$u = A\frac{xz}{r^3}, \tag{1a}$$

$$v = A\frac{yz}{r^3}, \tag{1b}$$

$$w = A\left(\frac{z^3}{r^3} + \frac{\lambda + 3\mu}{\lambda + \mu}\frac{1}{r}\right), \tag{1c}$$

where u, v, and w are displacement components in the x, y, and z directions, respectively, r is the distance from the origin, λ and μ are the Lamé constants for the material, and A is a scaling factor related to the applied force through:

$$A = \frac{F(\lambda + \mu)}{8\pi\mu(\lambda + 2\mu)}. \tag{2}$$

The Lamé constants can be related to the more commonly used modulus of elasticity and Poisson's ratio through:[16]

$$E = \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu}, \tag{3a}$$

$$\sigma = \frac{\lambda}{2(\lambda + \mu)}. \tag{3b}$$

Equation (1) predicts the displacement field resulting from the application of a force at a single point in space within an infinite uniform medium. However, acoustic radiation force is applied over the entire acoustic beam with a magnitude depending upon acoustic intensity, local attenuation, and local reflectivity. Thus Eq. (1), which is effectively a Green's Function, must be convolved spatially with the applied force field due to radiation force to yield the overall displacement field. Analytical solutions for the radiation force field in an attenuating medium are not generally available; however, by discretizing the problem and making assumptions which parallel those used to analyze acoustic heating,[17] it is possible to significantly reduce the complexity of the problem.

We begin by decomposing the forcing function into a series of disks spaced along the z-axis, with varying force magnitudes and radii. Within each disk we assume that the force exerted is constant. By spacing the disks closely along the z-axis, the overall forcing function can be accurately modeled as a series of disk forcing functions. The overall displacement field is then found by superimposing the displacement fields from each disk. While the three-dimensional displacement field would be of interest, we limit our analysis here to the displacement along the z-axis on the z-axis. This restriction allows an analytical solution of the response to a disk forcing function.

The displacement along the z-axis due to a disk forcing function centered at the origin with its axis of symmetry along the z-axis can be found by convolving the result of Eq. (1c) with a disk step function of radius a to yield:

$$w_{\text{disk}}(z) = \int_0^{2\pi}\int_0^a A\left(\frac{z^2}{(\sqrt{z^2 + r^2})^3} + \alpha\frac{1}{\sqrt{z^2 + r^2}}\right) r\, dr\, d\theta, \tag{4}$$

where we have changed to cylindrical coordinates and replaced $(\lambda + 3\mu)/(\lambda + \mu)$ by α to simplify notation. Carrying out the integrals yields:

$$w_{\text{disk}}(z) = 2\pi A\left((1 - \alpha)\sqrt{z^2} + \alpha\sqrt{z^2 + a^2} - \frac{z^2}{\sqrt{z^2 + a^2}}\right). \tag{5}$$

Equation (5) represents an analytical solution for the displacement field resulting from a disk forcing function with radius a centered at the origin. As stated earlier, the total displacement field can be found by summing the responses to a series of disk forcing functions:

$$w_{\text{tot}}(z) = \sum_{i=1}^{N} 2\pi A_i\left((1 - \alpha)\sqrt{(z - z_i)^2} + \alpha\sqrt{(z - z_i)^2 + a_i^2} - \frac{(z - z_i)^2}{\sqrt{(z - z_i)^2 + a_i^2}}\right), \tag{6}$$

where A_i is related to the force applied to disk i through Eq. (2), a_i is the radius of disk i, and z_i is the position of disk i along the z-axis. Equation (6) predicts the displacement along the z-axis resulting from a series of disk forcing functions spaced along the z-axis. This solution assumes an infinite uniform medium, and thus overestimates displacement by neglecting the confinement of tissue boundaries. This expression can be used to model the effects of acoustic radiation force by substituting in appropriate force magnitudes and disk radii resulting from acoustic radiation force. Throughout this paper the disk radii are determined using the analytical method described by Kossoff.[18] The radiation force exerted on each disk is given by[19]

$$F_i = \frac{W((z_{i-1} + z_i)/2) - W((z_i + z_{i+1})/2)}{c}, \tag{7}$$

where $W(z)$ is the acoustic power at range z, and c is the speed of sound. This expression calculates the lost acoustic power in the disk by taking the difference of the powers at the front and back of the disk. The acoustic power at any range can found by attenuating the source acoustic power:

$$W(z) = W_0 e^{-2\beta f z}, \tag{8}$$

where W_0 is the acoustic power applied by the transducer, β is the amplitude attenuation coefficient in Nepers per cm per MHz, and f is the frequency. The factor of 2 appears in this equation to account for the fact that attenuation coefficients are normally given in terms of amplitude, not intensity or power. This approach includes attenuation due to both scattering and absorption. Combining the effects of absorption and scattering introduces an error, as the radiation force induced by reflection is twice that resulting from absorption.[19] However, in soft tissues scattering only accounts for 10%–15% of attenuation,[20] causing the force to be underestimated by roughly 10%. Equation (8) could be easily expanded to allow attenuation which does not vary linearly with frequency or which varies as a function of range. It should be noted that higher forces may be generated at tissue boundaries where large specular reflection may occur.

THERMAL ANALYSIS

The application of acoustic radiation force in clinical settings will be limited by the risk of cavitation and acoustic

TABLE I. The acoustic and mechanical parameters assumed for the breast, liver, and vitreous.

Tissue	Attenuation	Acoustic velocity	Elastic modulus	Poisson's ratio
Breast	0.75 dB/cm/MHz (Ref. 23)	1510 m/s (Ref. 24)	1×10^4 N/m^2 (Ref. 25)	0.499 (Ref. 25)
Liver	1.17 dB/cm/MHz (Ref. 22)	1540 m/s (Ref. 22)	3.5×10^3 N/m^2 (bovine) (Ref. 26)	0.499 (Ref. 25)
Vitreous	0.10 db/cm/MHz (bovine) (Ref. 27)	1513.5 m/s (Ref. 28)	0.1 N/m^2 (Ref. 29)	0.499 (Ref. 25)

heating. While the likelihood of cavitation at a given acoustic intensity can be reduced by increasing the transmitted pulse length and reducing its amplitude, the heating associated with attenuation is more fundamental. A variety of techniques have been developed to predict acoustic heating in tissues.[17] Most begin with the Bioheat Transfer Equation, a partial differential equation describing heat dissipation in an environment which includes sources to account for acoustic heating and sinks to account for heat removal by perfusion. While these approaches are useful in many part of the body and yield analytical solutions for steady-state conditions, they can be tedious to employ when time dependent solutions in nonperfused tissues are desired.

In radiation force imaging the acoustic fields may be applied over a period of just a few seconds to tissues, such as the vitreous, which are not perfused by blood vessels. In these applications an analysis of heating which is time dependent and neglects perfusion is desirable. In other tissues, such as the breast and liver, which are well perfused, the short application times may make the effects of perfusion negligible. We begin by considering the time dependent solution for the temperature field along the z-axis resulting from a disk shaped heat source with radius a which has $\pi a^2 q\rho\gamma$ heat units instantaneously liberated over its face:[21]

$$T_\delta(z,t)=\frac{q}{2\sqrt{\pi\kappa t}}(1-e^{-a^2/(4\kappa t)})e^{-z^2/(4\kappa t)}, \tag{9}$$

where κ is the thermal diffusivity of the medium, ρ is the density, and γ is the heat capacity per unit mass. The derivation of acoustic heating presented here corrects an error in the derivation previously presented in Ref. 9. Radiation force imaging systems will not apply a single thermal impulse to the tissue, but will insonify tissue continuously over some specified time period. Since the temperature maximum should occur at the instant that the acoustic power is turned off, we can simplify the analysis by applying a thermal step function and evaluating the temperature field at the instant when the acoustic power is turned off. Equation (9) can be simplified somewhat by a substitution of variables to yield:

$$T_\delta(z,t)=\frac{q}{2\sqrt{\pi\kappa t}}(e^{-k_1^2/(4t)}-e^{-k_2^2/(4t)}), \tag{10}$$

where $k_1^2=z^2/\kappa$ and $k_2^2=(z^2+a^2)/\kappa$. By taking the Laplace transform of Eq. (10), multiplying by $1/s$ (the Laplace transform of a step function), and taking the inverse transform, it is possible to determine the temperature field resulting from heat deposited continuously in a disk beginning at time zero:

$$T_u(z,t)=\frac{q}{2\sqrt{k}}(2\sqrt{t/\pi}(e^{-m_1}-e^{-m_2})-(k_1\,\mathrm{erfc}(\sqrt{m_1}) - k_2\,\mathrm{erfc}(\sqrt{m_2}))), \tag{11}$$

where $m_1=k_1^2/(4t)$ and $m_2=k_2^2/(4t)$. Equation (11) predicts the temperature increase along the z-axis resulting from a disk shaped heat source placed at the origin and turned on at time zero. Since this expression now considers heat applied as a unit step function, $\pi a^2 q\rho\gamma$ is now equal to the heat deposited per second.

The result of Eq. (11) can be readily utilized within the framework for assessment of tissue displacement presented in this paper and the framework for assessing thermal risk presented in Ref. 17. Both approaches discretize the radiation field into a series of disks with various radii and acoustic intensities. The acoustic energy absorbed in each disk per unit time, which was also required in the displacement analysis, is equal to the heat applied to that disk per unit time. The temperature field from each of the disks are summed together to yield a total temperature distribution within the field. An equation like Eq. (6) can be readily expressed for the total temperature field.

EVALUATION AND RESULTS

The predicted beamwidth, temperature increase, and displacement due to acoustic radiation force was determined for vitreous, breast tissue, and liver. All cases assumed a 10-MHz piston transducer with a 0.5-cm-diameter aperture focused at a range of 1.0 cm. Acoustic fields were modeled for an acoustic intensity 1.0 W/cm^2 (in water) at the focus and an application time of 10 s. Note that the intensity described here and elsewhere refers to a uniform intensity over the beam area. The 10-s application time used here is a conservative estimate of the time required to form the entire radiation force image. By assuming that the formation of all image lines deposits heat along a single line, we produce a conservative estimate of temperature increase using a geometrically simple model. While the conditions described here may not be entirely realistic for imaging any of the tissues being examined, they do represent a reasonable synthesis of the requirements for imaging each of these tissues. Furthermore, by applying uniform conditions to each tissue it is possible to clearly depict the effects of tissue parameters. Displacement, beamwidth, and temperature increases were determined using the tissue parameters given in Table I.[22–29] In addition, tissue density was assumed to be 1.0 g/cm^3, thermal diffusivity was assumed to be 1.5×10^{-6} m^2/s,[30] and heat capacity was assumed to be 4×10^6 J/m^3 °C.[30] Param-

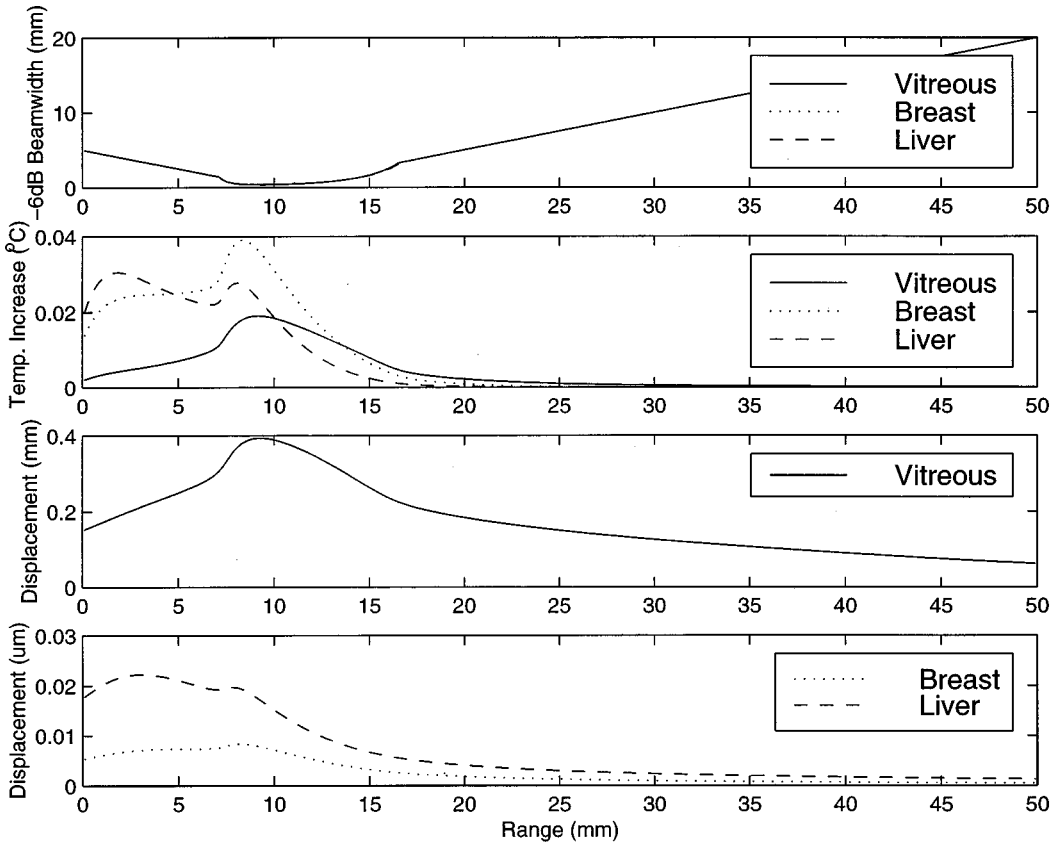


FIG. 1. The beamwidth predicted using the method described in Ref. 18, the thermal increase predicted by a superposition of disk sources using Eq. (11), and the displacement along the *z*-axis predicted by Eq. (6). The beamwidth plots cannot be differentiated because they overlap. The displacement of the vitreous is shown in the third panel, with that for the breast and liver shown in the fourth. Note the change in scale, indicating displacements in the breast and liver which are four orders of magnitude lower than that in the vitreous. All tissues were analyzed for a transducer with a 10-MHz operating frequency, a 0.5-cm diameter, and a focal range of 1.0 cm. The acoustic field was applied at an intensity of 1.0 W/cm^2 (in water) for 10 s.

eters obtained for human tissues were used when possible; in other cases bovine tissue parameters were used. The acoustic beam was modeled as a series of disks spaced 200 μm along the *z*-axis, with radius determined using the methods described in Ref. 18. These disks were used in conjunction with the methods presented in this paper to predict temperature increase and tissue displacement.

Figure 1 depicts the beamwidth predicted using the method described in Ref. 18, the thermal increase predicted by a superposition of disk sources using Eq. (11), and the displacement along the *z*-axis predicted by Eq. (6). The beamwidth plots cannot be differentiated because all overlap. The beamwidth near the transducer is determined by the transducer size. Beamwidth decreases to the diffraction limited resolution as the range increases toward the focus, then increases again as the range extends past the focus. All tissues were examined using the same transducer, so the only variation in beamwidth is due to the change in acoustic wavelength resulting from variations in the speed of sound between tissues.

The second set of curves in Fig. 1 depicts the temperature increase caused by a 10-s application of 10-MHz ultrasound using a piston transducer with an *f*-number of 2. These curves show a thermal increase in the breast which is twice that observed in the vitreous, and roughly 30% greater than that observed in the liver. The temperature profile in the vitreous has a strong maximum near the focus, with much less heating before and after the focus. The breast has an intermediate profile with a maximum at the focus, but significant heating before the focus. The liver shows another profile, with the global temperature maximum located near the transducer and a secondary peak located near the focus. These distributions can be readily understood when the attenuation of these tissues is considered. Because the liver has the highest attenuation of the tissues examined, it shows a significant temperature increase early in the propagation path. The vitreous has a very low attenuation coefficient, allowing the temperature increase to reach its maximum at the focus, where the acoustic intensity is at its maximum.

The third curve in Fig. 1 depicts the displacement field along the *z*-axis of the vitreous. The displacement is maximized near the focus, with a displacement of 400 μm. The fourth set of curves, which utilize a different scale, indicate the displacement field in the breast and liver. The displacement in these tissues is more than four orders of magnitude lower than in the vitreous. This tremendous variation in displacement results from the fact that the elasticity of breast and liver tissue is approximately four orders of magnitude greater than that of the vitreous.

The methods developed in this paper may be used to predict the radiation force induced displacement field for a variety of imaging systems and tissues. Specifically, the displacement field for a uniform intensity over the acoustic beam of 1.0 W/cm^2 (in water) was determined for a range of

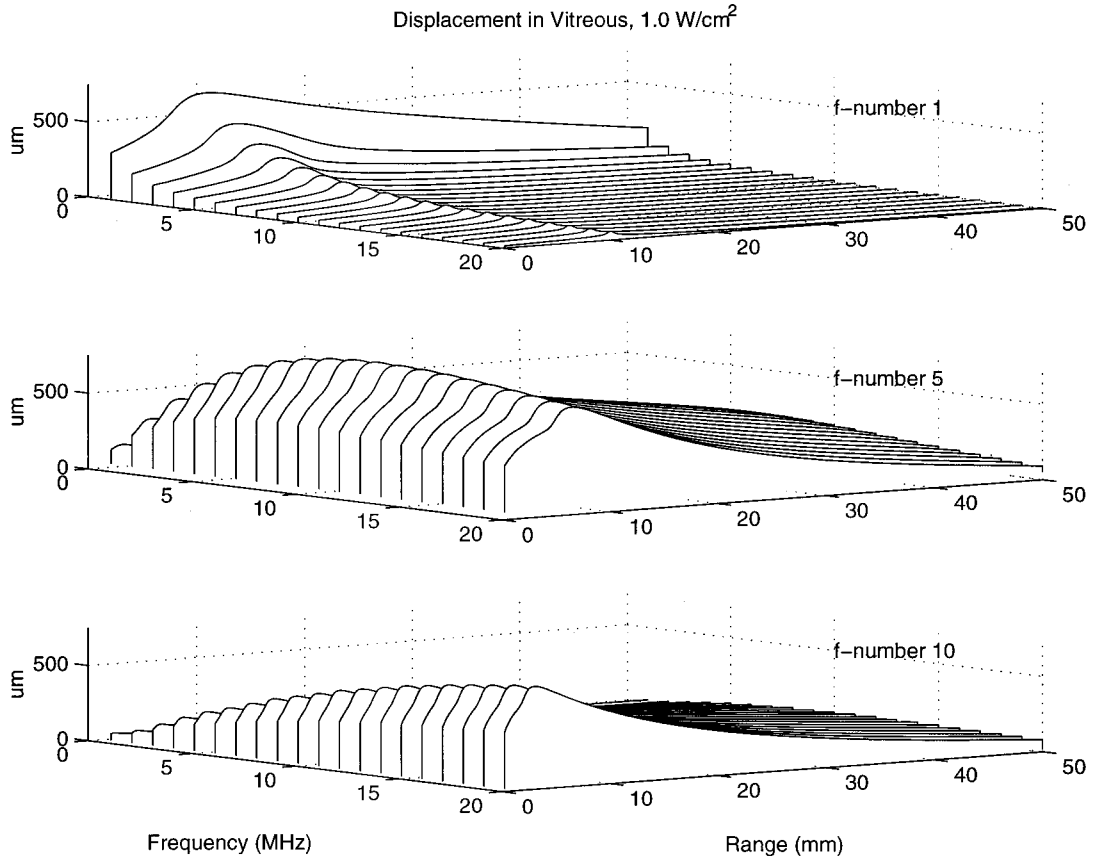


FIG. 2. Tissue displacement in the vitreous due to acoustic radiation force at a variety of *f*-numbers and operating frequencies. Each of the three graphs depicts the displacement as a function of range for operating frequencies between 1.0 and 20.0 MHz. The acoustic field was applied at an intensity of 1.0 W/cm^2 (in water) for 10 s. All transducers were focused at a range of 10 mm. The *f*-number was set at 1, 5, and 10 in the upper, middle, and lower graphs, respectively.

frequencies from 1.0 to 20.0 MHz and *f*-numbers of 1, 5, and 10. Note that the *f*-number is defined as focal range over transducer diameter. All transducers were focused at a range of 10 mm. Results for the vitreous are shown in Fig. 2 and results for the breast in Fig. 3.

Figure 2 depicts the displacement field induced in the vitreous for an acoustic field with a maximum intensity of 1.0 W/cm^2 (in water). In the upper figure the *f*-number was 1.0. This set of curves shows the displacement field becoming more compact, but decreasing in magnitude as the operating frequency increases. The peak displacement always occurs near the 10-mm focus of these *f*-number 1 imaging systems. In the second set of curves, corresponding to an *f*-number of 5, the displacement field still retains its maximum near the focal point, but the displacement shows a peak value at a frequency near 10.0 MHz. Finally, in the third set of curves, obtained for an *f*-number of 10, the peak displacement occurs between the transducer and the focus, with a maximum displacement occurring for a frequency of 20.0 MHz.

Figure 3 depicts the displacement field induced in the breast for an acoustic field with a maximum intensity of 1.0 W/cm^2 (in water). In each graph the operating frequency was varied between 1.0 and 20.0 MHz. In the upper graph, indicating results for an *f*-number 1 imaging system, the maximum displacement falls rapidly with increasing frequency, while the displacement field becomes more compact with increasing frequency. In the middle graph, corresponding to an *f*-number 5 system, the displacement field reaches a maximum between the focus and the transducer, and the displacement is maximized at roughly 5.0 MHz. In the lower graph, corresponding to an *f*-number of 10.0, the displacement falls off rapidly away from the transducer and reaches a maximum for a frequency of roughly 15.0 MHz. These results were qualitatively similar to those observed for liver, which are not shown here.

One goal in designing radiation force imaging systems is determining the conditions which generate the largest tissue displacement while maintaining the acoustic power within Food and Drug Administration (FDA) limits. To explore these situations the maximum displacement achievable in breast, vitreous, and liver tissue was determined using the methods described above. Operating frequency was varied between 1.0 and 20.0 MHz while the *f*-number was varied between 1 and 10, with transducers focused at a range of 10 mm. The acoustic intensity in the absence of attenuation was held at a constant value of 1.0 W/cm^2 for all situations. While this acoustic power level does not correspond directly to FDA limits in any particular tissue, it does allow easy conversion to more appropriate levels. The results of this analysis are shown in Fig. 4. In the vitreous, the displacement ranges up to nearly 1 mm, while in the breast and liver it peaks below 50 nm and 200 nm, respectively. In the vitreous the maximum displacement is generated over a broad range of conditions. In the breast and liver, however, the

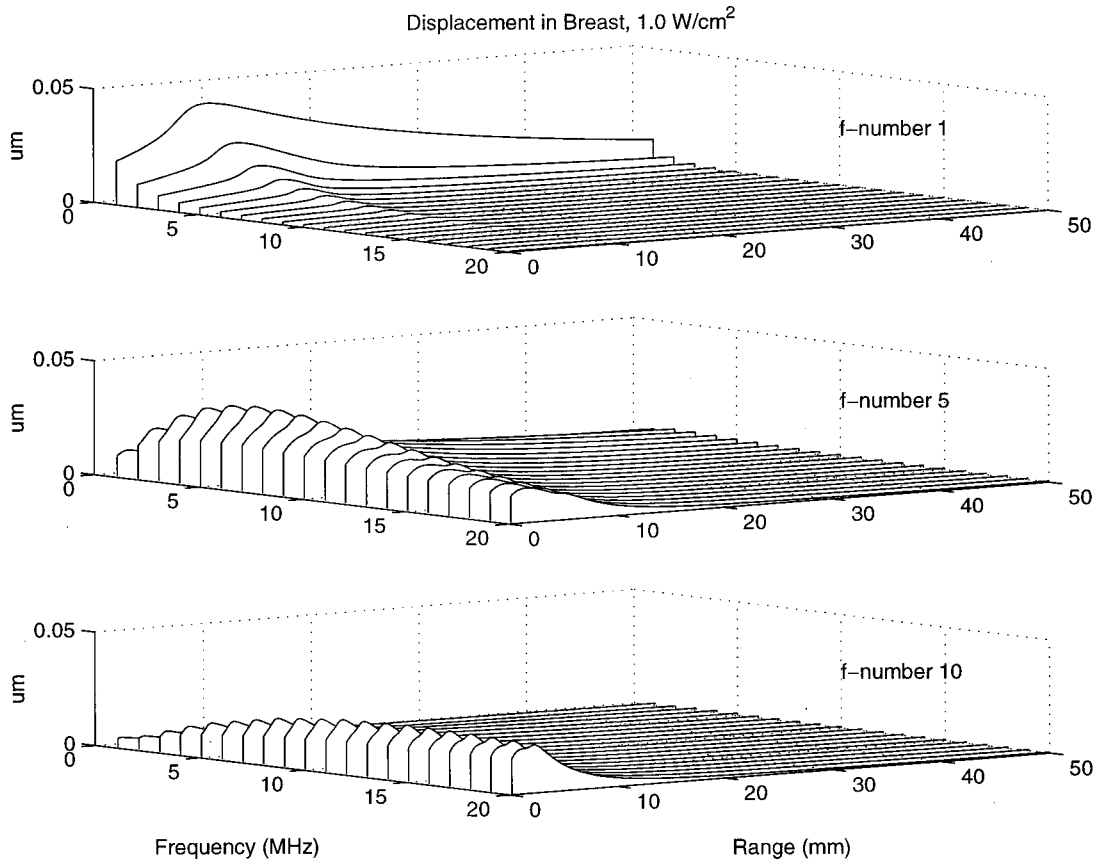


FIG. 3. Tissue displacement in the breast due to acoustic radiation force at a variety of *f*-numbers and operating frequencies. Each of the three graphs depicts the displacement as a function of range for operating frequencies between 1.0 and 20.0 MHz. The acoustic field was applied at an intensity of 1.0 W/cm^2 (in water) for 10 s. All transducers were focused at a range of 10 mm. The *f*-number was set at 1, 5, and 10 in the upper, middle, and lower graphs, respectively.

maximum displacements occur for low *f*-number and low frequency systems.

Another important question in the design of radiation force imaging systems is the uniformity, or alternatively the compactness, of the displacement field. This parameter was explored by determining the Full Width at Half-Maximum (FWHM) of the displacement field along the *z*-axis. As with the previous analysis, frequency was varied from 1.0 to 20.0 MHz and the *f*-number was varied from 1 to 10, with transducers focused at a range of 10 mm. Theoretical predictions are shown in Fig. 5. In the vitreous the maximum FWHM of the displacement field occurs over a broad range of frequencies and *f*-numbers. In the breast and liver, however, the FWHM is flat over most conditions and maximum for low *f*-numbers and frequencies.

Historical FDA limits on acoustic intensity are somewhat arbitrary, a more compelling limit is that imposed by the National Council on Radiation Protection and Measurements (NCRP), which states ''If the predicted maximum temperature elevation occurring in an ultrasound examination is less than 1 °C in an afebrile patient, there is no basis for believing that this exposure could lead to an adverse effect arising from thermal mechanism, regardless of exposure duration'' (p. 154 of Ref. 17). To explore the effects of this alternative constraint, the maximum displacement for a temperature increase of 1 °C was determined over a range of frequencies and *f*-numbers. For each condition the acoustic intensity was adjusted to elicit a temperature increase of 1 °C after a 10-s exposure. This exposure time is much greater than that needed to generate maximum displacement along a given direction. However, by considering the effects of long term interrogation along a single line we conservatively model the effects of the sequential, spatially distributed interrogations needed to form an image. Transducers were focused at a range of 10 mm. Results are shown in Fig. 6. In all tissues the maximum displacement occurs for very low frequencies and *f*-numbers. However, the breast and vitreous also show a secondary peak for low *f*-numbers and high frequencies. In the vitreous the maximum displacement is nearly 60 mm, while in the breast and liver it falls to 0.6 μm and 1.5 μm, respectively.

The validity of the model presented in this paper must be tested by comparison with other models and with experimental results. Sarvazyan *et al.* have developed models to predict the shear waves generated by pulsed and periodic radiation force under both linear and nonlinear conditions.[12] In one test case they consider radiation force applied by a 3.0-MHz ultrasound transducer with a focus at 5.0 cm and a radius of 2.0 cm. The ultrasound energy was modulated at a frequency of 1.0 KHz and had an intensity at the transducer face of 10 W/cm^2. It should be noted that these conditions would result in an acoustic intensity at the focus of approximately 1.3 KW/cm^2. While this intensity far exceeds that which is normal in diagnostic ultrasound, it will not necessarily damage tissue if applied for only a few ms as the authors propose. The tissue was modeled with a density of 1

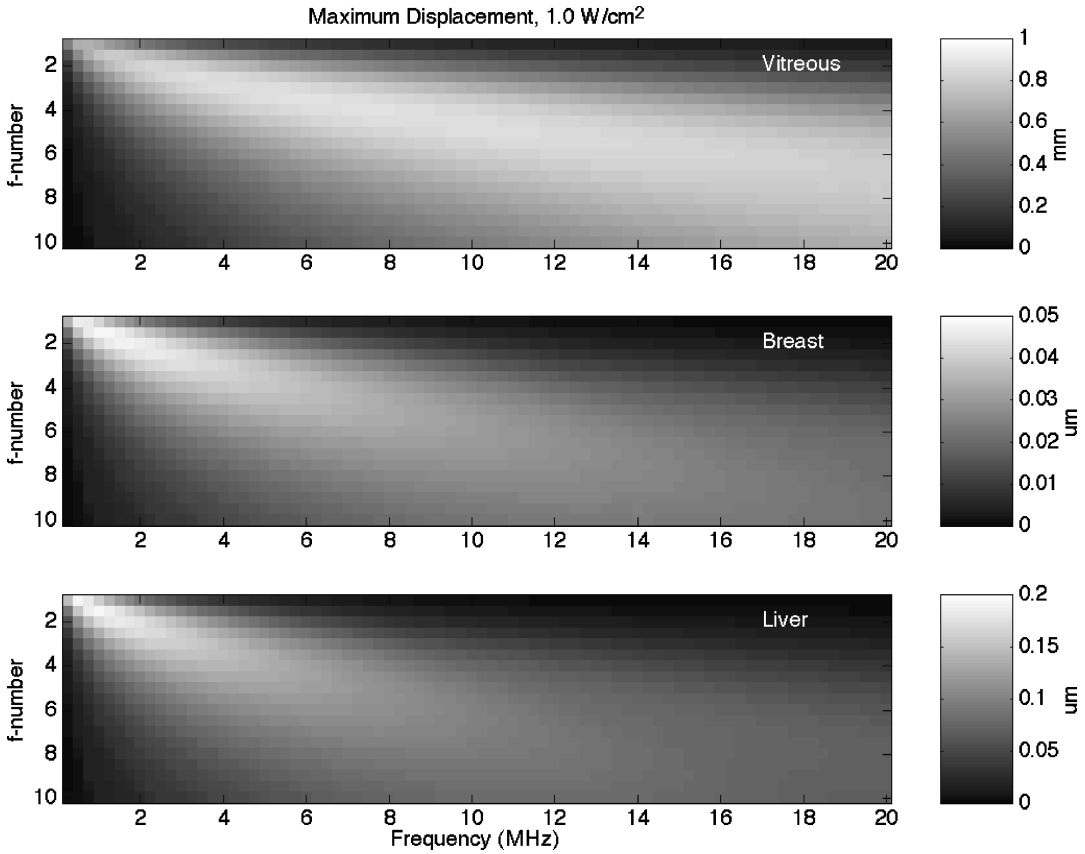


FIG. 4. The maximum tissue displacement due to acoustic radiation force in vitreous, breast, and liver tissue. Operating frequency and *f*-number (focal range over transducer diameter) were varied to determine their effects on the applied radiation force. The acoustic field was applied at an intensity of 1.0 W/cm^2 (in water) for 10 s, with a focal range of 10 mm. Note that the scale of the displacement varies for each tissue examined.

g/cm^3, a speed of sound of 1500 m/s, and a modulus of elasticity of 2.7×10^4 N/m^2 (27 kPa). These conditions yielded a maximum displacement of approximately 35 μm. Applying these conditions to the model presented here (neglecting modulation and assuming an attenuation coefficient of 1.17 dB/cm/MHz) yielded a maximum displacement of 16.5 μm. While the maximum displacements predicted by the two models agree within a factor of 2, the axial displacement profile differs dramatically. Sarvazyan's model shows a displacement field which is symmetrical and very localized around the focus. The model presented here predicts significant displacement along the entire propagation path from the transducer to the focus. The displacement maximum, 16.5 μm, occurs at a range 1.0 cm from the transducer face. The displacement falls to roughly 10.0 μm at the 5.0-cm focus. One possible error in comparison was in the selection of the attenuation coefficient. Sarvazyan states that the attenuation coefficient for liver was used in the model. While Ref. 22 indicates a value of 1.17 dB/cm/MHz, a value of 0.5 dB/cm/MHz is often assumed. The model of this paper was tested again using this value with the above conditions. The maximum displacement then occurred at the 5.0-cm focus, but had fallen to 14.8 μm. In addition a plateau in displacement occurred between 2.0 and 4.0 cm, with a displacement of 13.5 μm. Both models predict maximum displacements on the same order of magnitude, but the shape of the displacement field differs dramatically.

Andreev *et al.* have performed experiments utilizing optical methods to estimate the displacements induced in tissue mimicking materials by acoustic radiation force.[31] Experiments were performed using two transducers with center frequencies of 0.96 and 1.8 MHz and acoustic power outputs of 70 and 45 W, respectively. Both transducers had a focus at 70 mm and a diameter of 64 mm. A $20\times20\times23$ mm cube of tissue mimicking material was interrogated. This material exhibited a speed of sound of 980 m/s, a density of 0.98 g/cm^3, a modulus of elasticity of 1.38×10^4 N/m^2, and an attenuation coefficient of 0.06 cm^{-1} at 0.96 MHz and 0.09 cm^{-1} at 1.8 MHz. The tissue mimicking block was placed at the transducer focus within a water filled tank. The conditions described here yield an acoustic intensity of 3.6 KW/cm^2 at 0.96 MHz and 8.2 KW/cm^2 at 1.8 MHz. As with earlier examples, these intensities may not result in significant thermal risk if application times are extremely short. These conditions were modeled using the methods described in this paper. The attenuation coefficient was varied axially to account for water in the propagation path. The first 6.0-cm region had an attenuation of 0.05 dB/cm/MHz to model water; the next 2.0-cm region used the attenuation described above for the tissue mimicking material; and finally a third domain utilized an attenuation of 0.05 dB/cm/MHz. In the experiments described by Andreev, both transducers yielded a maximum displacement of approximately 20 μm. In the analysis presented here the predicted displacement was 32.1 μm at 0.96 MHz and 40.4 μm at 1.8 MHz. In both cases the agreement between theory and experiments is good. Discrep-

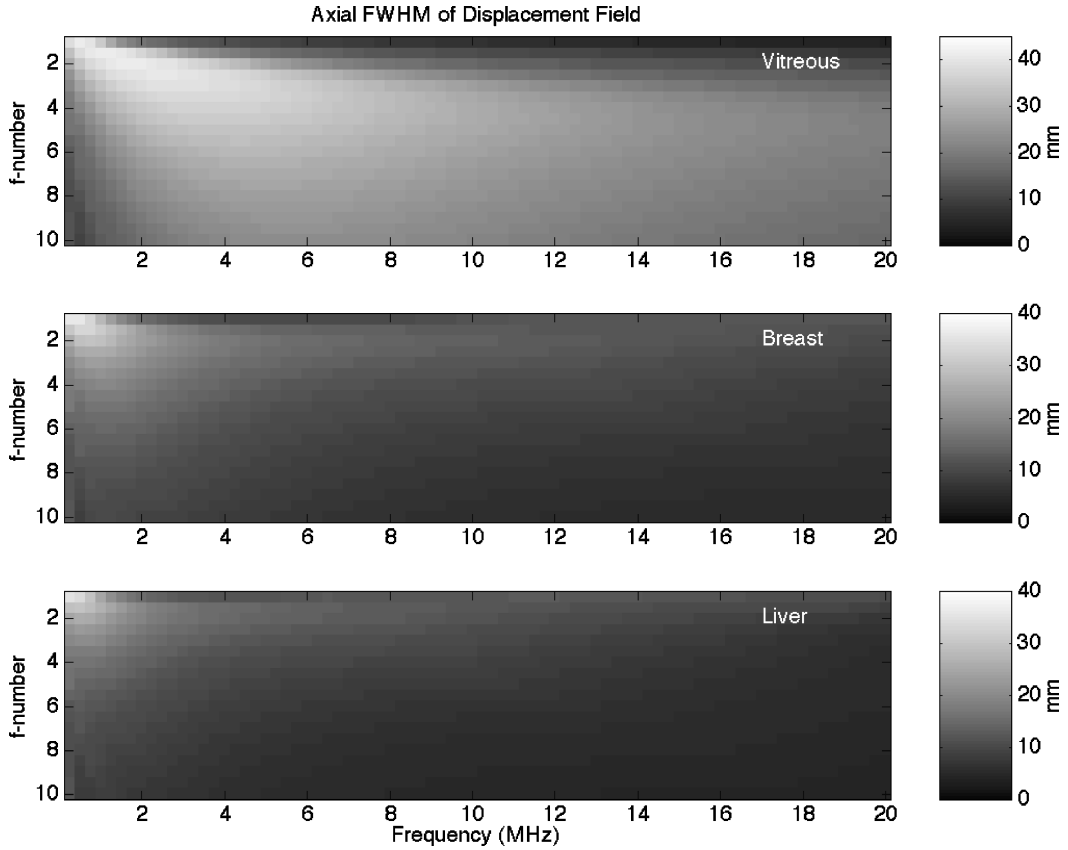


FIG. 5. The axial Full Width at Half Maximum (FWHM) of the displacement field induced in vitreous, breast, and liver tissues. Operating frequency and *f*-number (focal range over transducer diameter) were varied to determine their effects on the FWHM of the displacement field. The acoustic field was applied at an intensity of 1.0 W/cm^2 (in water) for 10 s, with a focal range of 10 mm.

ancies may result because water regions in the model were treated as an elastic solid which would transmit force to the tissue mimicking material, while water in the experiment was of course a fluid.

The model presented in this paper can also be compared to the results of experiments performed by the author.[9] These experiments examined the vitreous body of porcine eyes which had been softened, to better mimic aged human vitreous, by the injection of 500 units of Hyaluronidase. Eyes were insonified at 5.0 MHz using eight cycle pulses with a pulse repetition frequency (PRF) of 12.0 KHz. Experiments utilized a piston transducer with a focal length of 50.8 mm and a diameter of 9.5 mm. These conditions were modeled using tissue parameters given in Table I and transducer geometry identical to that described for the experiments. The acoustic power output for these experiments was not measured, making a direct comparison with theory impossible. However, based on other similar experiments in which the acoustic power was measured, we expect an acoustic intensity on the order of 5.0 mW/cm^2 at the focus yielding a predicted displacement of 15.3 μm. This is in good agreement with the measured displacements of 5–40 μm. While this comparison cannot fully confirm the accuracy of the presented model, it does indicate reasonable validity.

DISCUSSION

The analysis presented here compares favorably to the experimental work of the author,[9] with the analysis predicting displacements on the order of 15 μm, and experiments yielding displacements between 5 and 40 μm at an assumed acoustic intensity of 5.0 mW/cm^2. Such a displacement should be readily detectible using motion tracking methods including cross correlation or pulsed Doppler. Also, these displacements are generated at an acoustic power which falls below the FDA historical acoustic power limit in the eye of 68 mW/cm^2 (in water).[32] Thus radiation force imaging in the vitreous has the potential to be both successful and safe.

Good agreement was also found between the predictions of the theory presented here and experimental and theoretical work by other groups.[12,31] Under conditions with acoustic intensities on the order of KW/cm^2, all methods found displacements on the order of tens of microns for materials which mimic breast or liver tissue. These displacements may be reasonably measured, although these acoustic intensities could not be easily generated with current diagnostic ultrasound equipment. The risks associated with ultrasonic heating would be minimal under these conditions as long as application times of only a few ms were utilized. Further work is needed, however, to explore the possibility of damage due to cavitation.

The method of analysis presented here is only one of a variety of possible approaches. The acoustic beam pattern could be readily modeled using the analytical approach developed in Refs. 33 and 34, or by a variety of numerical simulation methods. The beam pattern could be used as an input to the methods presented here for computation of the

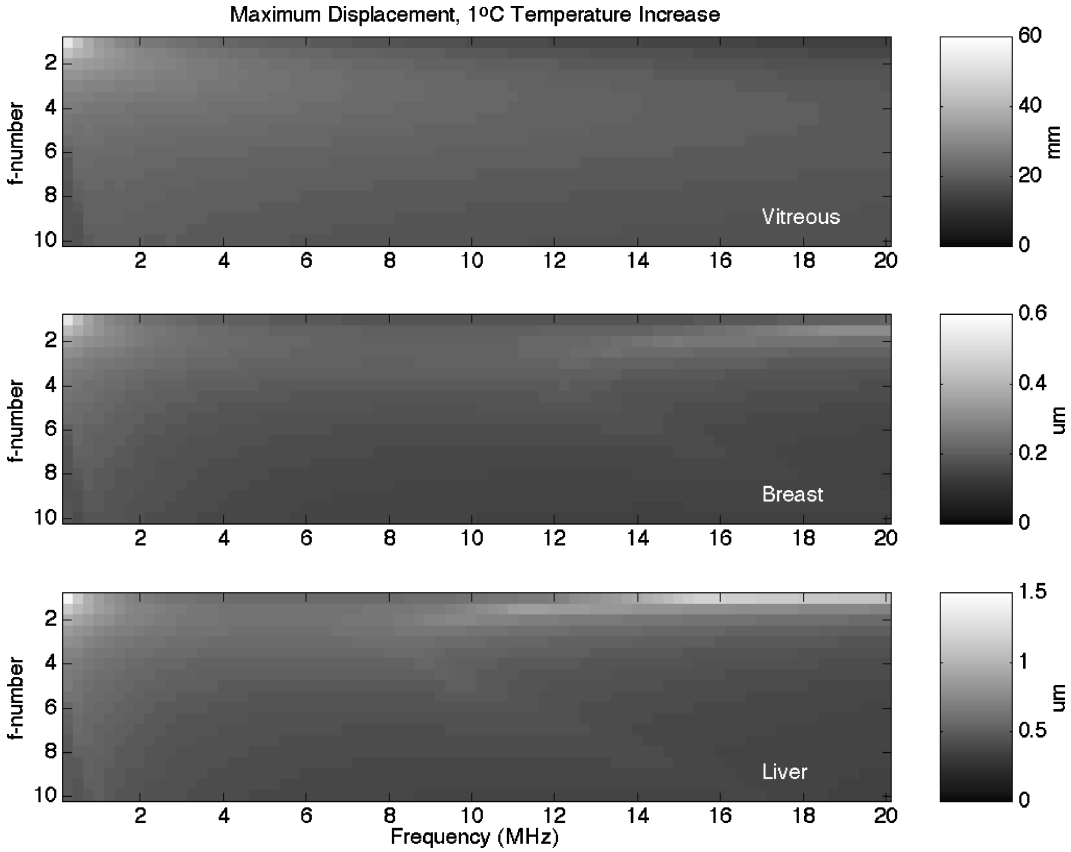


FIG. 6. The maximum tissue displacement due to acoustic radiation force in vitreous, breast, and liver tissue for a 10-s acoustic application with an intensity adjusted to yield a 1 °C temperature increase. Operating frequency and *f*-number (focal range over transducer diameter) were varied to determine their effects on the applied radiation force. All transducers were focused at a range of 10 mm. Note that the scale of displacement varies for each tissue examined.

forcing function under linear conditions, or those presented in Ref. 34 under either linear or nonlinear conditions. With the spatial forcing function defined, the disk method described here could be used, or a direct convolution with the displacement Green's function could be performed. Of course these modifications would still retain the assumption of an infinite uniform medium. This assumption could be eliminated through use of a more sophisticated Green's function which includes two or more tissue regions. Finally, the most general analysis could be performed by combining acoustic simulation tools with Finite Element Methods to determine displacements. While that approach is attractive, it would be extremely time consuming to analyze the broad range of conditions considered in this paper.

One factor which has been considered only superficially in this paper is the attenuation and resultant heating due to variations in tissue type in the acoustic propagation path. In the eye, for example, the acoustic beam must pass through a water path, the sclera, and possibly through the lens before reaching the vitreous. On the opposite side of the eye, the exiting acoustic wave passes through the retina, sclera, orbital fat, and finally the bony orbit. All of the tissues mentioned have attenuation coefficients which are significantly higher than that of the vitreous. Thus these tissues will exhibit relatively high thermal increases, especially if high *f*-number systems are utilized. We have performed initial simulations which indicate that the temperature increases can be reduced by nearly 50% by avoiding insonification of the lens.[9] Further analysis will be needed to determine the appropriate transducer configuration and operating frequency for radiation force imaging in realistic, complex tissues.

The method of analysis presented in this paper considers insonification at only a single frequency. This approach is valuable for many proposed radiation force imaging systems which would utilize narrow-band insonification. In systems which utilize broadband insonification the methods presented here could be used to analyze the displacement generated at a number of frequencies, with the total displacement found through the application of superposition.

Acoustic heating of tissue may cause a variety of confounding effects for an imaging system based on acoustic radiation force. First, the expansion of tissue with acoustic heating will cause tissue displacement independent of that due to radiation force. The thermal expansion coefficient of soft tissues is approximately $3\times10^{-4}/°\mathrm{C}$,[20] meaning that a 1 °C temperature increase will cause tissue to expand to 1.0003 times its original volume. Assuming that the tissue is free to expand uniformly in all dimensions, this would yield and axial expansion of 1.0001 times. For a block of tissue 2.0 cm on a side, this would cause the axial length to change by approximately 2.0 μm. This exceedingly small displacement is relatively large when compared to the radiation force induced displacements predicted in this paper. Another artifact results from the apparent displacement which would be observed when the speed of sound in tissue changes due to a temperature increase. A 1 °C temperature increase causes a

speed of sound increase of approximately 2 m/s/°C.[20] For our hypothetical 2.0-cm block of tissue, the far wall would appear to move 26 μm, an amount which exceeds most of the displacements predicted in this paper. To limit artifacts, radiation force should be generated with minimal temperature increase. A temperature increase of only 1 °C would introduce artifacts which would overwhelm the displacements induced by radiation force in either the liver or breast. In the vitreous these limitations may be less pronounced because of the larger displacements generated, the more uniform nature of the tissue, and a lack of perfusion which would create a nonuniform temperature field.

Two different strategies may be employed in radiation force imaging. In the first, a relatively uniform radiation force field is applied and variations in tissue properties are indicated by variations in the induced displacement field. This approach would allow generation of relatively large displacements, but may be susceptible to artifacts and exhibit poor resolution because an inhomogeneity would disrupt the entire displacement field. In the second approach a compact force field could be applied sequentially to a variety of locations, with tissue variations indicated by variations in the local displacement. This approach may yield better spatial resolution, but will probably achieve lower maximum displacements for a given acoustic intensity, as indicated by the results of Fig. 2. Comparison of these two methods will require a more sophisticated analysis than the one presented in this paper. Specifically, the analysis must include the effects of inhomogeneities in tissue elasticity.

Regardless of the strategy chosen to generate radiation force, received echoes must be processed to estimate induced displacements. This estimation is subject to errors because of finite signal bandwidth, finite center frequency, a limited window of available data, the presence of electronic noise, and signal decorrelation caused by physical distortions of the tissue. The errors in displacement estimation caused by these factors can be quantified using the Cramér–Rao Lower Bound.[35] Assuming a 10.0-MHz center frequency, an 80% bandwidth, a signal window length of 1.0 mm (1.3 μs), no signal decorrelation, and root mean square (rms) signal to noise ratio of 30 yields a predicted rms error in displacement estimation of only 0.12 μm. If the signal decorrelation is increased to 0.9, a very significant level of tissue distortion, the rms error in displacement estimation increases to 1.3 μm. In either case the displacement estimation error is relatively small when compared to the potential errors induced by thermal effects. Thus while displacement estimation errors cannot be completely disregarded, their significance seems to be minimal.

The analysis in this paper indicates that a perfectly uniform displacement field cannot be obtained for the transducer geometries examined. One approach to improving displacement field uniformity would be to superimpose a number of acoustic beams with varying focal depths and operating frequencies, like the method proposed by Hassock to improve transmit depth of field in B-Mode imaging.[36]

The methods of predicting displacement presented in this paper assume that the tissue is a purely elastic material. However, it is well known that biological tissues exhibit viscoelastic behavior. In experiments performed in the vitreous[9] this property appeared as a creep of the displacement over time. In one experiment the vitreous was displaced a distance of 40 μm over a period of approximately 0.13 s. At the end of the observation period the vitreous was still moving a significant amount. Imaging of tissue viscoelastic properties offers an exciting application for radiation force imaging. However, this property also restricts the minimum observation time needed to achieve accurate measures of tissue parameters. The observer must apply radiation force long enough that the tissue is displaced enough for accurate measurement. There may be some inherent tradeoff between the application time required and the maximum force exerted, with increasing forces applied to reduce interrogation time. There will also be a tradeoff between the rate of thermal increase and required application time.

The analysis presented in this paper neglects the impact of nonlinear acoustic propagation on the displacement field. Nonlinear propagation effectively shifts some of the applied acoustic energy into higher frequencies which are more rapidly attenuated and thus generate higher than expected acoustic radiation forces. The analysis of displacement due to radiation force presented in this paper could be coupled with the analysis of radiation force due to nonlinear propagation[34] to yield a more comprehensive picture of the displacement due to nonlinear effects.

CONCLUSION

Analytical and numerical methods were developed to predict the tissue displacement due to acoustic radiation force and the temperature increases due to attenuation in nonperfused tissues. In the breast and liver the maximum displacements were on the order of hundreds of nm for acoustic intensities of 1.0 W/cm^2. In the vitreous, however, because it is four orders of magnitude softer, the maximum achievable displacements were on the order of hundreds of microns. These results indicate that radiation force imaging may be easily applied to the vitreous, but may be less practical for stiffer tissues such as the breast and liver.

ACKNOWLEDGMENTS

I would like to acknowledge a series of helpful discussions with Katherine R. Nightingale. This work was supported in part by NIH SBIR Grants Nos. R43-EY11337 and R43-EY11456.

[1] M. O'Donnell, S. Y. Emelianov, A. R. Skovoroda, M. A. Lubinski, and B. M. Shapo, ''Quantitative elasticity imaging,'' Proc. IEEE Ultrason. Symp., pp. 893–903 (1993).
[2] J. Ophir, I. Céspedes, H. Ponnekanti, Y. Yazdi, and X. Li, ''Elastogrphy: A quantitative method for imaging the elasticity of biological tissues,'' Ultrason. Imaging **13**, 111–134 (1991).
[3] Y. Yamakoshi, J. Sato, and T. Sato, ''Ultrasonic imaging of internal vibration of soft tissue under forced vibration,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **37**, 45–53 (1990).
[4] R. M. Lerner, K. J. Parker, J. Holen, R. Gramiak, and R. C. Waag, ''Sono-elasticity: Medical elasticity images derived from ultrasound signals in mechanically vibrated targets,'' Acoust. Imaging **16**, 317–327 (1988).
[5] W. F. Walker, B. H. Friemel, L. N. Bohs, and G. E. Trahey, ''Real-time imaging of tissue vibration using a two-dimensional speckle tracking system,'' Proc. IEEE Ultrason. Symp., pp. 873–877 (1993).

[6] R. Muthupillai, D. J. Lomas, P. J. Rossman, J. F. Greenleaf, A. Manduca, and R. L. Ehman, ''Magnetic resonance elastography by direct visualization of propagating acoustic strain waves,'' Science **269**, 1854–1857 (1995).
[7] T. Sugimoto, S. Ueha, and K. Itoh, ''Tissue hardness measurement using the radiation force of focused ultrasound,'' Proc. IEEE Ultrason. Symp., pp. 1377–1380 (1990).
[8] M. Fatemi and J. F. Greenleaf, ''C-scan imaging by radiation force stimulated acoustic emission method,'' Proc. IEEE Ultrason. Symp. **2**, 1459–1462 (1996).
[9] W. F. Walker, R. E. Davidsen, and C. A. Toth, ''Applications of acoustic radiation force in ophthalmic ultrasound,'' Proc. IEEE Ultrason. Symp. **2**, 1291–1295 (1997).
[10] W. F. Walker, R. E. Davidsen, and C. A. Toth, ''A novel method of ocular kinetic ultrasound examination,'' presented at 22nd International Symposium on Ultrasonic Imaging and Tissue Characterization, Rosslyn, VA, abstract published in Ultrason. Imaging **19** (1997).
[11] S. Y. Emelianov, S. D. Swanson, J. B. Fowlkes, O. V. Rudenko, and A. P. Sarvazyan, ''Remote surrogate palpation: shear wave elasticity imaging,'' presented at 22nd International Symposium on Ultrasonic Imaging and Tissue Characterization, Rosslyn, VA, abstract published in Ultrason. Imaging **19** (1997).
[12] A. P. Sarvazyan, O. V. Rudenko, S. D. Swanson, J. B. Fowlkes, and S. Y. Emelianov, ''Shear wave elasticity imaging—A new ultrasonic technology of medical diagnostics,'' Ultrasound Med. Biol. **24**, 1419–1436 (1998).
[13] K. R. Nightingale, ''Ultrasonic generation and detection of acoustic streaming to differentiate between fluid-filled and solid lesions in the breast,'' Ph. D. Thesis, Duke University (1997).
[14] K. R. Nightingale, P. J. Kornguth, W. F. Walker, B. A. McDermott, and G. E. Trahey, ''A novel ultrasonic technique for differentiating cysts from solid lesions: Preliminary results in the breast,'' Ultrasound Med. Biol. **21**, 745–751 (1995).
[15] A. E. H. Love, *A Treatise on the Mathematical Theory of Elasticity* (Dover, New York, 1944).
[16] G. S. Kino, *Acoustic Waves* (Prentice Hall, Englewood Cliffs, 1987).
[17] *Exposure Criteria for Medical Diagnostic Ultrasound: I. Criteria Based on Thermal Mechanisms*, Vol. NCRP Report No. 113 (National Council on Radiation Protection and Measurement, Bethesda, MD, 1992).
[18] G. Kossoff, ''Analysis of focusing action of spherically curved transducers,'' Ultrasound Med. Biol. **5**, 359–365 (1979).
[19] P. N. T. Wells, *Biomedical Ultrasonics* (Academic, London, 1977).
[20] F. A. Duck, *Physical Properties of Tissue: A Comprehensive Reference Book* (Academic, San Diego, 1990).
[21] H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids* (Clarendon, Oxford, 1947).
[22] R. A. Mountford and P. N. T. Wells, ''Ultrasonic liver scanning: The quantitative analysis of the normal A-scan,'' Phys. Med. Biol. **17**, 14–25 (1972).
[23] F. S. Foster and J. W. Hunt, ''Transmission of ultrasound beams through human tissue-focusing and attenuation studies,'' Ultrasound Med. Biol. **5**, 257–268 (1979).
[24] G. Kossoff, E. K. Fry, and J. Jellins, ''Average velocity of ultrasound in the human female breast,'' J. Acoust. Soc. Am. **53**, 1730–1736 (1973).
[25] A. P. Sarvazyan, A. R. Skovaroda, S. Y. Emelianov, J. B. Fowlkes, J. G. Pipe, R. S. Adler, R. B. Buxton, and P. L. Carson, ''Biophysical bases of elasticity imaging,'' Acoust. Imaging **21**, 223–240 (1995).
[26] S. R. Huang, ''Principles of sonoelasticity imaging and its applications in hard tumor detection,'' Ph. D. Thesis, University of Rochester (1990).
[27] L. Filipczynski, J. Etienne, G. Lypacewicz, and J. Salkowski, ''Visualizing internal structures of the eye by means of ultrasonics,'' Proc. Vib. Probl. Warsaw **8**, 357–368 (1967).
[28] C. L. De Korte, A. F. W. Van Der Steen, and J. M. Thijssen, ''Acoustic velocity and attenuation of eye tissues at 20 MHz,'' Ultrasound Med. Biol. **20**, 471–480 (1994).
[29] R. L. Zimmerman, ''*In vivo* measurements of the viscoelasticity of the human vitreous humor,'' Biophys. J. **29**, 539–544 (1980).
[30] K. M. Sekins and A. F. Emery, ''Thermal science for physical medicine,'' in *Therapeutic Heat and Cold*, edited by J. F. Lehmann (Williams and Wilkins, Baltimore, MD, 1982), pp. 70–132.
[31] V. G. Andreev, V. N. Dmitriev, Y. A. Pishchalnikov, O. V. Rudenko, O. A. Sapozhnikov, and A. P. Sarvazyan, ''Observation of shear waves excited by focused ultrasound in a rubberlike medium,'' Acoust. Phys. **43**, 149–155 (1995).
[32] ''Information for Manufacturers Seeking Marketing Clearance of Diagnostic Ultrasound Systems and Transducers,'' U.S. Department of Health and Human Services; Food and Drug Administration, Center for Devices and Radiological Health, September 30, 1997.
[33] M. B. Vinogradova, O. V. Rudenko, and A. P. Sukhorukov, *Theory of Waves*, 2nd ed. (Nauka, Moscow, 1990).
[34] O. V. Rudenko, A. P. Sarvazyan, and S. Y. Emelianov, ''Acoustic radiation force and streaming induced by focused nonlinear ultrasound in a dissipative medium,'' J. Acoust. Soc. Am. **99**, 2791–2798 (1996).
[35] W. F. Walker and G. E. Trahey, ''A fundamental limit on delay estimation using partially correlated speckle signals,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42**, 301–308 (1995).
[36] J. A. Hossack, ''Extended focal depth imaging for medical ultrasound,'' Proc. IEEE Ultrason. Symp. **2**, 1535–1540 (1996).

The potential of transskull ultrasound therapy and surgery using the maximum available skull surface area

Jie Sun and Kullervo Hynynen
Division of MRI/Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115

(Received 11 September 1998; revised 27 November 1998; accepted 16 January 1999)

Based on previous studies,[1–3] the transskull ultrasound field is re-examined by utilizing the maximum available skull surface area. The source is assumed to be in direct contact with the skull outer surface, and phase correction is adopted to obtain a sharp focus at a desired location. A digitized skull profile was obtained from Magnetic Resonance (MR) scan images of a volunteer. Two driving frequencies (0.5 and 1.0 MHz) within the appropriate frequency range for transskull ultrasound therapy and surgery are investigated. With no phase correction, there is no apparent transskull focus. With phase correction, a sharp transskull focus is obtained at the desired location. Both pressure and specific absorption rate (SAR) gains (ratios of pressure amplitude and SAR at the focal point compared to those on the outer skull surface) are examined, and it is shown that the skull heating problem can be overcome by utilizing the maximum available skull surface area. By specifying the phase correction for different locations, the focus can be successfully moved inside the deep brain volume without significantly compromising the pressure and SAR gains; however, the sidelobes may be of concern at superficial sites. © *1999 Acoustical Society of America.* [S0001-4966(99)03404-9]

PACS numbers: 43.80.Sh, 43.80.Vj [FD]

INTRODUCTION

Recently, there has been renewed interest in focused ultrasound therapy and surgery (see the review articles by ter Haar and Hill[4,5]). One of the potential applications of focused ultrasound is for brain therapy and surgery. Since the early experiments of Fry *et al.*,[6–10] it has been known that the human skull strongly attenuates and distorts ultrasound-beam propagation, and that transskull ultrasound therapy and surgery may be possible but difficult. Smith *et al.* introduced the phased-array technique to ultrasound diagnostic imaging in order to correct the beam degradation due to the presence of intervening bone.[11,12] Thomas and Fink proposed to use this technique in therapeutic ultrasound,[13] and the results of time-reversal focusing and steering through the skull using a one-dimensional small-element array have been reported.[14] On the other hand, *in vivo* bioeffects of ultrasound in the animal brain have been studied by many investigators,[15–18] and most recently, Vykhodtseva *et al.* have shown the appropriate exposure conditions to cause predictable thermal damage and to separate different types of bioeffects in the brain tissue.[19,20]

In our previous studies,[1–3] we have shown the feasibility of using focused ultrasound for brain treatments through an intact human skull by using the phased-array technology. By using a classical spherically curved phased array with frequencies below 1.0 MHz, it is not only possible to generate a sharply focused transskull ultrasound beam, but also to overcome the skull overheating problem.[3] However, by examining the pressure and the specific absorption rate (SAR, which is a good indicator of tissue temperature elevation) at the focal point compared to those on the skull surfaces, it is clear that only marginal pressure and SAR gains have been achieved, particularly at frequencies above 1.0 MHz. A numerical model has been developed which can use digitized layer interfaces to calculate ultrasound wave absorption, diffraction, reflection, and refraction.[2,3,21–23] A similar approach, which used full elastic wave equations, has recently been reported separately by Kühnicke.[24–27] He showed an example in nondestructive testing (NDT) to calculate the acoustic field from a foil transducer through a steel shaft, and an example in medical ultrasound to calculate the acoustic field through a schematic eye. It has been verified experimentally that the phase shifts needed to correct the wavefront distortion due to the presence of an intervening human skull can be calculated from outer- and inner-skull surfaces derived from medical images.[2]

In the present study, the transskull ultrasound field will be re-examined by utilizing the maximum available skull surface area, in order to increase the pressure and SAR gains. The source is assumed to be in direct contact with the outer-skull surface which covers the entire skull segment, and phase correction is applied to each of the simple sources in order to obtain a sharp focus at a desired location inside the brain. The purpose of this configuration is to estimate the best performance limitations for the transskull ultrasound field, in terms of achievable pressure and SAR gains, the sidelobes, and the region in which the focus can be successfully moved.

I. THEORY

A. A 2-layer problem

Consider a source that is in direct contact with the outer-skull surface; the basic problem of interest is to evaluate the acoustic pressure field in the brain due to the specified normal velocity distribution on the outer-skull surface. This is a straightforward 2-layer propagation problem. The skull and the brain are denoted as layer 1 and 2, the outer-skull surface

 0001-4966/99/105(4)/2519/9/$15.00 © 1999 Acoustical Society of America

is denoted as interface 1, and the inner-skull surface is denoted as interface 2. The acoustic pressure in the brain can be readily evaluated using the following formula,[3]

$$p(x,y,z)=\sum_{i_2=1}^{N_2}\frac{jk_{c_2}\rho_2 c_2}{2\pi}\frac{e^{-jk_{c_2}R}}{R}(u_{i_2}\,ds_{i2})$$
$$=\sum_{i_1=1}^{N_1}\sum_{i_2=1}^{N_2}\left[\frac{jk_{c_1}}{2\pi}\frac{e^{-jk_{c_1}R_{i_1,i_2}}}{R_{i_1,i_2}}\times\left(1-j\frac{1}{k_{c_1}R_{i_1,i_2}}\right)T_{v_{i_2}}\cos(\theta_{i_2}^t)\right]$$
$$\times\left[\frac{jk_{c_2}\rho_2 c_2}{2\pi}\frac{e^{-jk_{c_2}R}}{R}\right]u_{i_1}(ds_{i_1}\,ds_{i_2}). \qquad (1)$$

The skull surfaces were divided into small patches, and ds_{i_1} and ds_{i_2} are used to denote the areas of small patches on the outer- and inner-skull surfaces, respectively, with u_{i_1} and u_{i_2} being the corresponding normal velocities. If the dimensions of ds_{i_1} and ds_{i_2} are small enough compared with the acoustic wavelength, a simple source can be defined for a small patch[28] with its source strength being denoted as $(u_{i_1}\,ds_{i_1})$ or $(u_{i_2}\,ds_{i_2})$. Clearly, ρ_1, c_1 are the density and sound speed of the skull and ρ_2, c_2 are the density and sound speed of the brain. $k_{c_1}=k-j\alpha_1$, $k_{c_2}=k-j\alpha_2$ are the complex wave numbers of the first and second layer, respectively, with α_1 and α_2 being the corresponding attenuation coefficients. $(x_{i_1},y_{i_1},z_{i_1})$ denotes the location of a simple source $(u_{i_1}ds_{i_1})$ on the outer skull surface, $(x_{i_2},y_{i_2},z_{i_2})$ denotes the location of a simple source $(u_{i_2}ds_{i_2})$ on the inner-skull surface, and (x,y,z) denotes the location of a field point in the brain. $R_{i_1,i_2}=\sqrt{(x_{i_1}-x_{i_2})^2+(y_{i_1}-y_{i_2})^2+(z_{i_1}-z_{i_2})^2}$ and $R=\sqrt{(x-x_{i_2})^2+(y-y_{i_2})^2+(z-z_{i_2})^2}$. It is assumed that there are total N_1 simple sources on the outer-skull surface, and N_2 simple sources on the inner-skull surface. On the inner-skull surface, the transmission coefficient for the particle velocity can be written as,

$$T_{v_{i_2}}=\frac{2[\rho_1 c_1/\cos(\theta_{i_2}^i)]}{\rho_2 c_2/\cos(\theta_{i_2}^t)+\rho_1 c_1/\cos(\theta_{i_2}^i)}\frac{\cos(\theta_{i_2}^i)}{\cos(\theta_{i_2}^t)}, \qquad (2)$$

where the incidence and transmission angles $\theta_{i_2}^i, \theta_{i_2}^t$ satisfy the Snell's law, i.e., $\sin\theta_{i_2}^t/\sin\theta_{i_2}^i=c_2/c_1$.

B. Digitized interface

For real human anatomy, e.g., a human skull, the layer interface will come in the form of digitized data profiles. The data from a medical imaging modality, e.g., CT or MRI, can be cross-sectional images which are obtained with a fixed slice thickness.

As illustrated in Fig. 1, a very simple approach has been developed for surface reconstruction using cross-section contours, i.e., for two adjacent contour lines, (1) lateral resampling/interpolation is performed so that two adjacent contour lines contain an equal number of nodes, and all the lateral distance between adjacent nodes should be shorter than half of the acoustic wavelength; (2) if any longitudinal gaps between the two adjacent contour lines are greater than half of the acoustic wavelength, longitudinal resampling/interpolation has to be performed, i.e., extra contour lines have to be added in between; (3) a quadrangle is now defined by four nodes that are adjacent to each other; and (4) assuming the interface is continuous and so is its first-order derivative, each small piece represented by four nodes will be treated as a simple source.

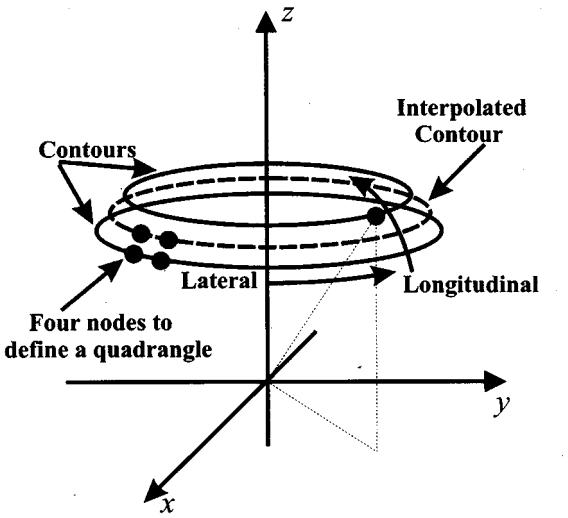


FIG. 1. Resampling/interpolation for two contour lines.

With several points in space, there are many classical formulas readily available to calculate the area and the corresponding normal direction, e.g., with three points in space, a plane can be defined along with its normal direction, and the triangle area can be evaluated with known coordinates of these three points.

C. Phase correction

With the source in direct contact with the outer-skull surface, it is clear that without phase correction, there won't be any distinguishable transskull foci inside the skull, because of the skull's irregular shape, nonuniform skull thickness, and very high sound speed in the skull. The purpose of phase correction is to introduce phase offset to source elements so that the acoustic wave radiated from each source element will add up coherently at a desired focal point. The optimum case is that the radiated acoustic wave from each source element arrives at the desired focal point with the same phase. In the present study, the phase correction will be applied to each of the simple sources on the outer-skull surface.

In order to obtain a phase offset for each of the simple sources on the outer-skull surface, the reverse problem is considered.[3] In the reverse problem, a point source is radiating at the desired focal point, and the complex pressure at each simple source location on the outer-skull surface can be evaluated all at once, which will then be fed into the forward problem to evaluate the acoustic field inside the brain using Eq. (1).

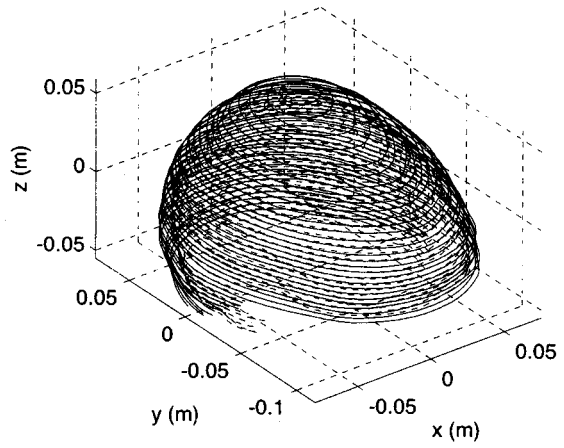


FIG. 2. A digitized human skull profile.

D. SAR

The equation relating the specific absorption rate (SAR) and the pressure amplitude is presented here for future reference,

$$\mathrm{SAR}=\alpha\frac{|p|^2}{\rho c}, \tag{3}$$

where $|p|$ is the pressure amplitude, α is the pressure attenuation coefficient, ρ and c are density and sound speed of the medium, respectively. Clearly, SAR is proportional to the pressure amplitude square.

II. SIMULATIONS

A digitized human skull profile was constructed from MR scan images. Figure 2 shows the 3-D skull traces at different MRI slices for the skull's outer and inner surfaces, respectively. The coordinate origin has been chosen at a selected center inside the skull, which will simply be called the skull center in the rest of the paper.

Table I shows the layer properties. The acoustic properties of the human skull used in this paper are mainly derived according to a paper published by Fry and Barger.[8] The density and the sound speed of the skull were obtained by weighted averages combining inner table, dipole, and outer table. Assuming the source is in direct contact with the outer-skull surface, it is clearly a two-layer propagation problem.

Consider the acoustic waves originated from the skull center. They will propagate out towards the inner-skull surface first; then on the inner-skull surface, part of the waves will be reflected back and part of the waves will penetrate the skull bones. If the mean distance from the skull center to the skull inner surface is around 7.5 cm and the typical normal incidence (maximum) brain–skull velocity reflection coefficient is around 1/3, a rough estimation shows that up to 25% and 15% of the waves could be reflected back to the skull center at 0.5 and 1.0 MHz, respectively, assuming all the reflected waves could arrive at the skull center in phase. In reality, due to the skull's irregular shape, not all of the skull reflected waves will add up coherently at any points inside the skull. It is shown in Fig. 3 that the maximum pressure amplitude obtained without phase correction [in (a) and (c)] is less than 10% of the focal amplitude obtained with phase correction [(b) and (d)]. This means that the maximum value of the reflected waves would be less than 2.5% of the focal amplitude, which is insignificant. Hence, the skull reflections after the beam has propagated through the center can be ignored for the purpose of this study. The simulation using Eq. (1) did not take into account the multiple reflections on the inner-skull surface.

It is desirable to examine its resulting ultrasound field in the brain, in particular, to compare the peak acoustic pressure at the focal point to those on the outer-skull surface (It is known that the maximum pressure amplitude of the whole skull volume should appear on the outer skull surface.[3]) It is also desirable to examine the extent of a region in which the focus could be moved by aligning the phases of the simple sources on the outer-skull surface for a particular location in the brain.

At 0.5 and 1.0 MHz, Fig. 3 shows the acoustic pressure in the x-y plane centering at (0, 0, 0) cm without and with phase correction. In order to have a better view of the acoustic-field pattern, both surface and contour plots are presented, and the contour lines are drawn with the interval of 10% of the peak value. For both frequencies, without phase correction, there is no focus inside the brain; with phase correction, a clear focus has been observed at (0, 0, 0) cm. Due to the asymmetry of the skull, it is necessary to at least also examine the acoustic field in x-z and y-z planes. At 0.5 and 1.0 MHz, Fig. 4 shows the acoustic pressure in both x-z and y-z planes centering at (0, 0, 0) cm with phase correction, which further confirms that desirable transskull acoustic-field distribution can be achieved. The phase-corrected cross-sectional beam profiles of these acoustic fields are also plotted in Fig. 5 along x, y, and z axes. The classical 3-dB intensity roll-off lines are illustrated so that the cross-sectional beamwidths can be roughly estimated.

TABLE I. The properties of 2 layers: the skull–brain layers in front of the source.

			A 2-layer skull–brain model								
			Attenuation (nep/m)								
			Frequency (MHz)								
	Density (kg/m^3)	Sound speed (m/s)	0.5	0.625	0.75	0.875	1.0	1.125	1.25	1.375	1.5
Skull[8]	1796	2652	50	50	85	137	179	223	315	390	464
Brain	1030	1545	4.0× frequency								

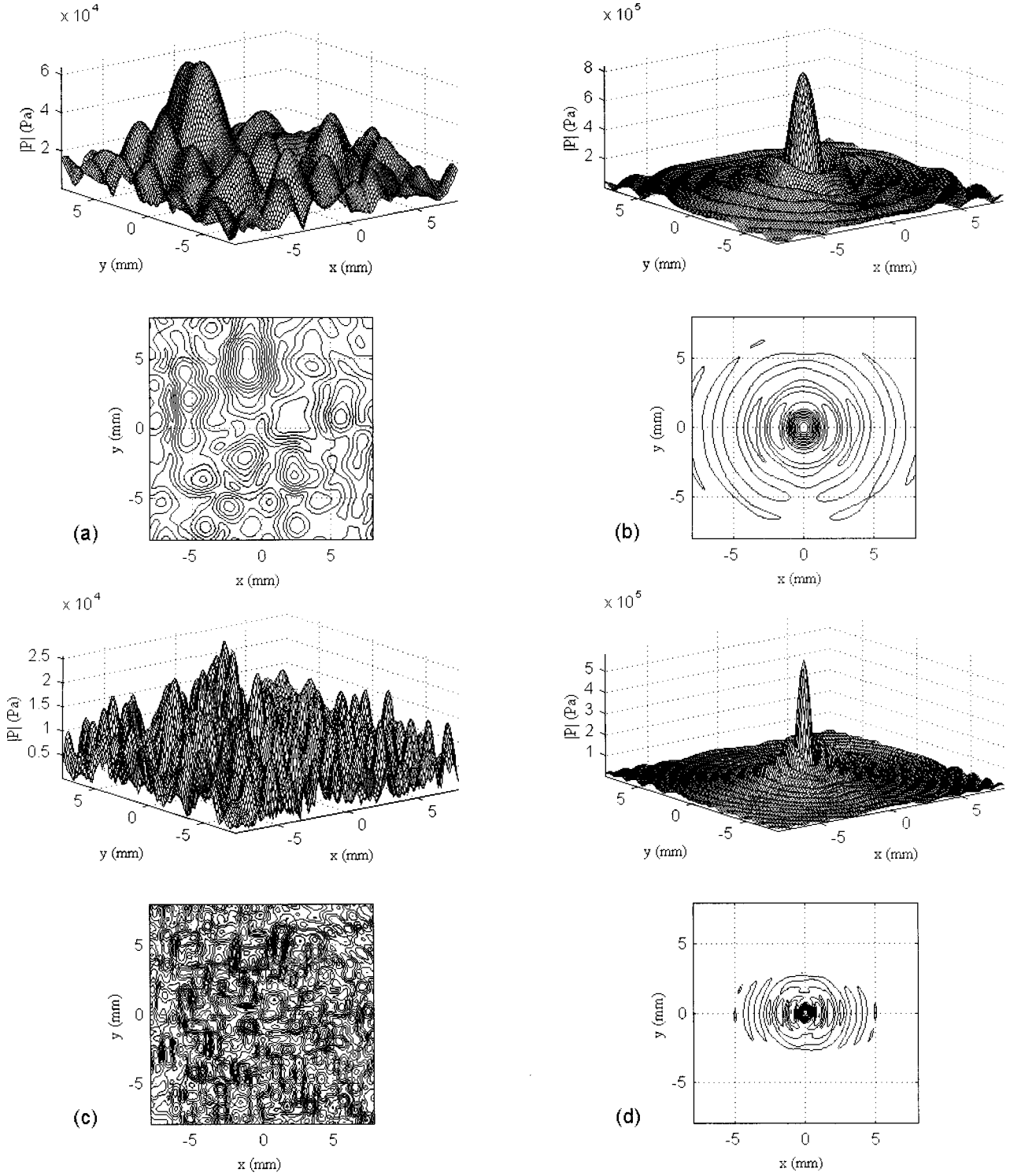


FIG. 3. The transskull ultrasound field in the x-y plane when the focus is at the center: (a) and (c) without phase correction, (b) and (d) with phase correction; (a) and (b) at 0.5 MHz, (c) and (d) at 1.0 MHz.

Beamwidths of less than half of the wavelength have been observed. Figure 6 is 3-D visualization of contour surfaces for the focal volume inside the brain at 0.5 and 1.0 MHz; the contour surfaces represent the locations for 30% [(a) and (c)], 50% (6-dB intensity roll off), 70% (~3-dB intensity roll off), and 90% [(b) and (d)] of the peak acoustic-pressure amplitude. Sidelobes are observed only at the 30% level. As expected, the focal volume dimensions at 1.0 MHz are roughly half of those at 0.5 MHz.

Although all the ultrasound fields presented in the paper were within the vicinity of the focal point, extended spatial windows were used to examine the ultrasound fields within the whole brain volume to make sure there were no significant sidelobes or grating lobes elsewhere. It was observed, for focusing at the skull center, that there was no noticeable energy concentration outside the vicinity of the focal point; on the other side, for focusing off the skull center, particularly at a shallow location, there was noticeable pressure

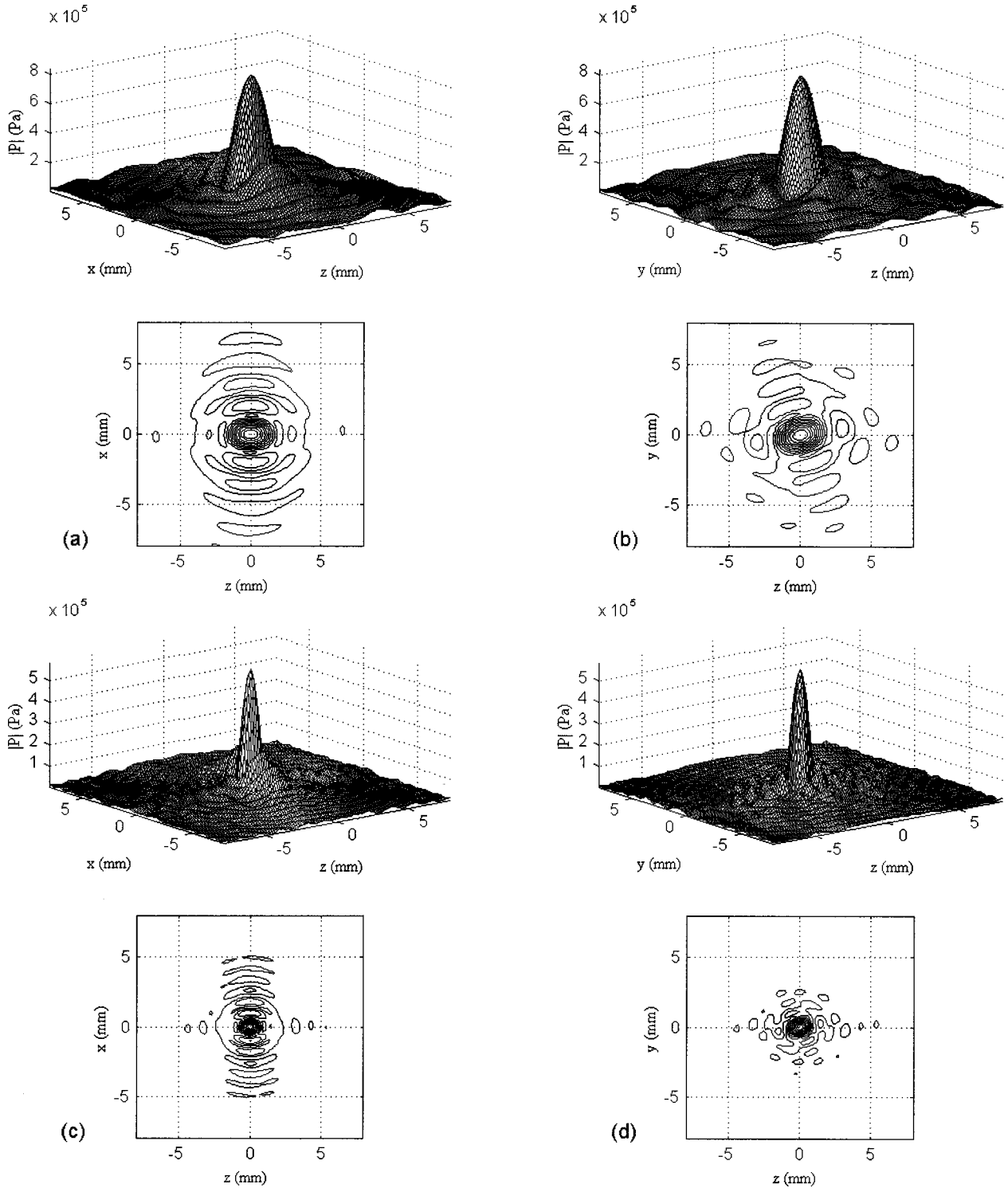


FIG. 4. The transskull ultrasound field in the x-z and y-z planes when the focus is at the center with phase correction: (a) and (b) at 0.5 MHz, (c) and (d) at 1.0 MHz.

maxima outside the vicinity of the focal point. For focusing close to the skull surface, these pressure maxima could be up to 20% of the focal amplitude. Nevertheless, these pressure maxima (if there were any) were generally of much lower levels compared to the focal amplitude and the sidelobe levels observed in the vicinity of the focal point.

A span of frequencies from 0.25 to 1.25 MHz has been considered to examine the pressure and SAR gains, in order to investigate skull heating after utilizing the maximum available skull surface area. Figure 7 shows both the pressure and SAR gains as a function of frequency. Both the pressure and SAR gain curves peak at 0.60–0.65 MHz. Considering only the pressure, Fig. 8 shows the percentage of the highest sidelobe amplitude in the x-y plane compared to the focal amplitude vs frequency, and the lowest is also observed at 0.60–0.65 MHz. It is observed in many simulations that, for an arbitrarily located focal point inside the skull, the sidelobes along any direction, including those in the x-y, x-z,

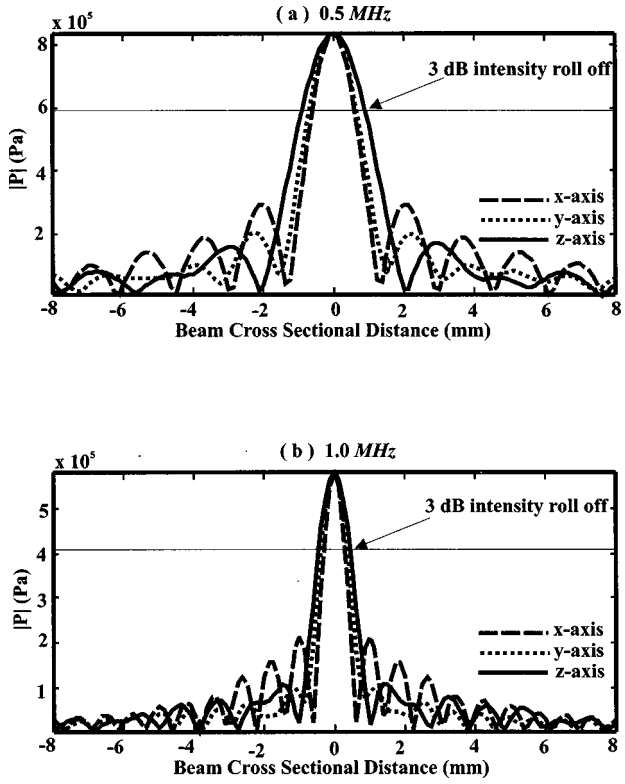


FIG. 5. Cross-sectional beam profiles along *x*, *y*, and *z* axes: (a) at 0.5 MHz, (b) at 1.0 MHz.

and *y*-*z* planes, are of comparable amplitudes. Hence, in this paper, the highest sidelobe amplitude in the *x*-*y* plane is chosen as a representative to be compared with the focal amplitude. Although this approach is not exhaustive, it does provide a reasonably good estimation. Figure 9 shows the change of the 3-dB intensity roll-off beamwidth along the *x*, *y*, and *z* axes with respect to the acoustic wavelength. Table II shows the peak acoustic-pressure amplitude at the focal point and on the outer-skull surface. The focal dimensions along the *x*, *y*, and *z* axes have also been estimated. These

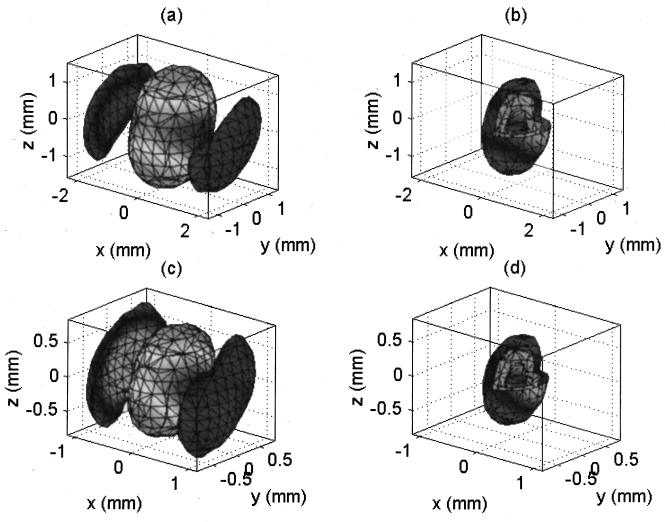


FIG. 6. Contour surfaces for the focal volume when the focus is at the center: (c) and (d) at 1.0 MHz, (a) and (b) at 0.5 MHz; (a) and (c) are 30% contour surfaces, (b) and (d) are 50%, 70%, and 90% contour surfaces.

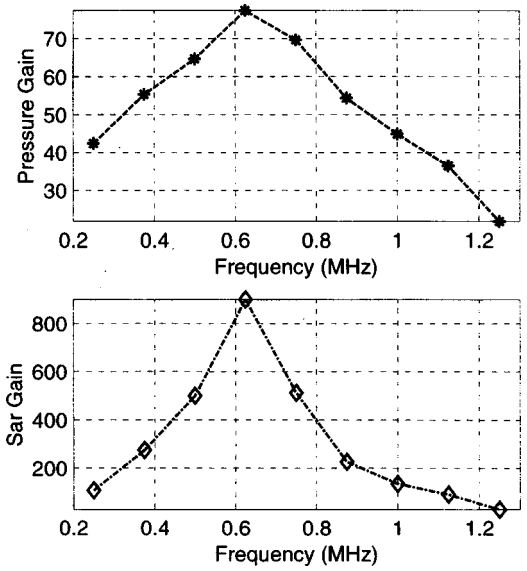


FIG. 7. Pressure and SAR gains versus frequency when the focus is at the center.

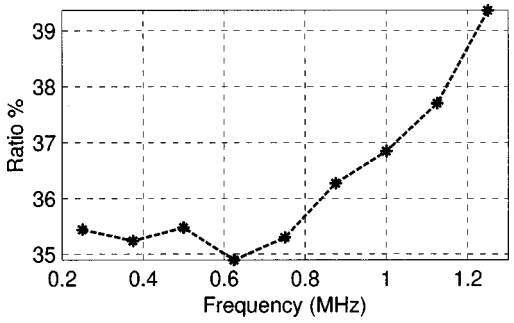


FIG. 8. The percentage of the highest sidelobe pressure amplitude in the *x*-*y* plane compared to that at the focal point versus frequency when the focus is at the center.

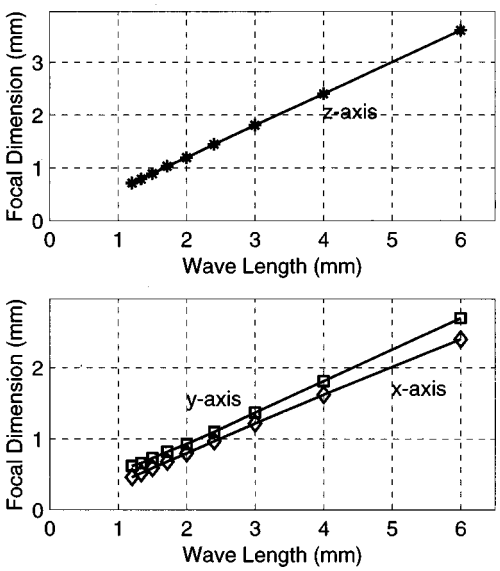


FIG. 9. The 3-dB intensity roll-off beamwidth along *x*, *y*, and *z* axes versus wavelength when the focus is at the center.

TABLE II. The comparison of the peak acoustic-pressure amplitudes, pressure gains, and focal dimensions.

		Peak pressure amplitude ($Pa\times10^3$)			Focal dimensions (mm)	
	Frequency (MHz)	Focal point	Outer skull surface	Pressure gain ζ_0	D_t x/y axis	D_a
Previous[3] (16×16)	0.5	232	60	3.9	3.2/3.2	20.3
	1.0	160	61	2.6	1.6/1.6	10.2
Current (phase-corrected)	0.5	836	13	64	1.2/1.4	1.80
	1.0	582	13	44	0.6/0.7	0.89

numbers are compared with the earlier results for the spherical cap transducer (F number is 1.0).[3] By examining the current pressure and SAR gains, a significant increase has been observed.

In order to find a region inside the skull in which the focus can be successfully moved, it is necessary to examine the focused ultrasound field when the focal point is being moved along transverse, vertical, and diagonal directions, due to the asymmetry of the skull. At 0.5 and 1.0 MHz, Figs. 10, 11, and 12 show both the pressure and SAR gains when the focus inside the brain has been moved off the skull center in transverse, vertical, and diagonal directions, respectively. The limits for moving along these selected directions are around 4.0, 2.0, and 2.0 cm off the center, before you will see significant drop-off of pressure and SAR gains. The gains decrease when the focus is being moved off the center; at the same time, the peak amplitudes of sidelobes are noticeably increased. Considering only the pressure, the change of the percentage of the highest sidelobe amplitude in the x-y plane compared to the focal amplitude is shown in Fig. 13, when the focus is being moved off the center in three different directions, and Fig. 14 shows the contour lines for the pressure amplitude in the x-y plane when the focus has been moved diagonally to locations of (1.5, 1.5, 1.5) and (3.0, 3.0, 3.0) cm.

III. DISCUSSION

The aim of this paper was to explore the best-case scenario for transskull ultrasound therapy and surgery when the whole skull surfaces were used for ultrasound transmission. By examining the maximum achievable pressure and SAR gains when the whole skull surfaces are utilized, it is shown that the skull overheating problem can be eliminated. The optimum frequency for transskull ultrasound transmission is around 0.6–0.65 MHz, which confirms our previous study.[3] On the other hand, considering moving the focus inside the skull, it is shown that the focus can be moved successfully within a deep volume of less than $8.0\times8.0\times4.0\,\mathrm{cm}^3$ without significantly compromising the pressure and SAR gains. However, as the focus is moving away from the selected center, the focal amplitude is decreasing, and the sidelobes are growing. For therapeutic purposes, the focus can be moved away from the center further in a transverse direction than in a vertical direction. Multiple foci appeared at superficial target locations, and their locations around the focal point are not controllable.

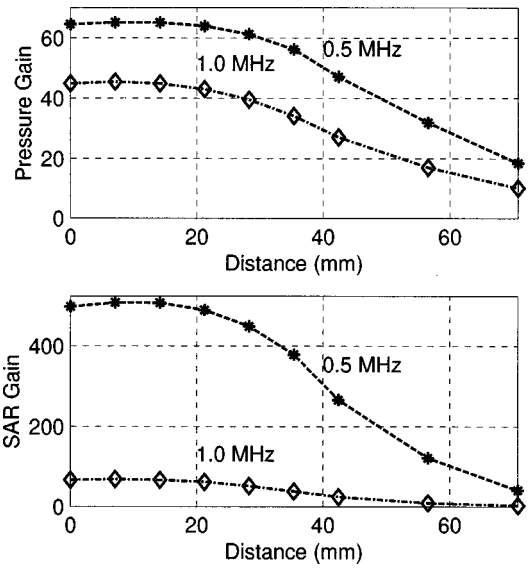


FIG. 10. Pressure and SAR gains versus the distance when the focal point is being moved away from the center in a horizontal direction at both 0.5 and 1.0 MHz.

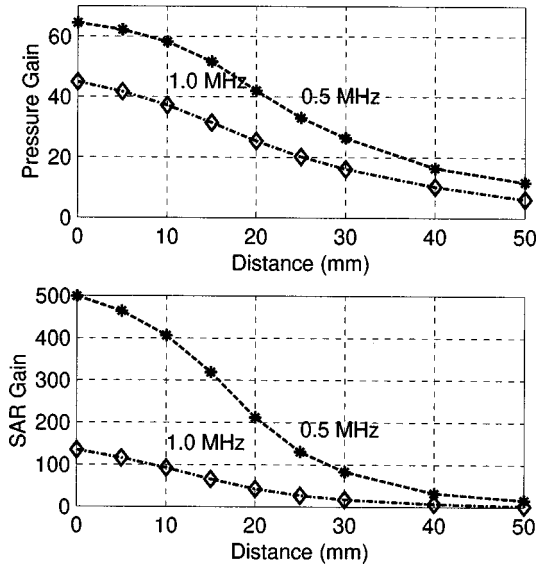


FIG. 11. Pressure and SAR gains versus the distance when the focal point is being moved away from the center in a vertical direction at both 0.5 and 1.0 MHz.

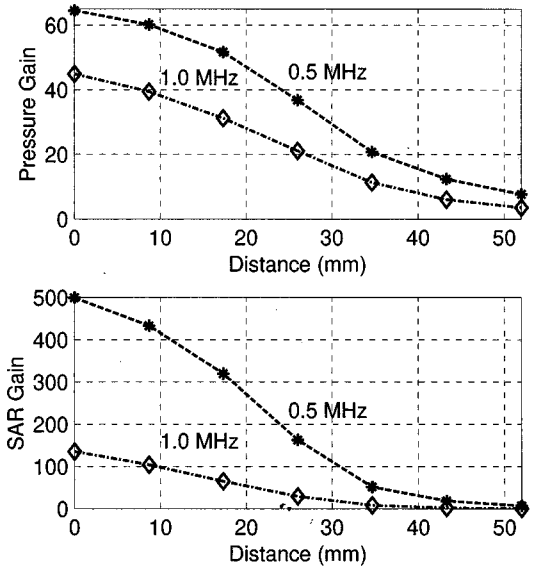


FIG. 12. Pressure and SAR gains versus the distance when the focal point is being moved away from the center in a diagonal direction at both 0.5 and 1.0 MHz.

For the classical spherical cap transducer studied before,[3] the penetration area on the outer skull surface is less than 90 cm^2, while in the current study, the utilized outer-skull surface area is around 565 cm^2, a factor of more than sixfold change of the pressure gain can be attributed to the increase of the outer-skull surface penetration (radiating) area, i.e., the focal amplitude is approximately proportional to the skull penetration area. Uniform pressure distribution on the skull surfaces is another key to eliminate the skull heating problem. In the current study, the source is specified to be a uniform velocity distribution which is in direct contact with the outer-skull surface. By examining the peak versus average pressure amplitude on the outer-skull surface for the case previously studied,[3] a factor of around a threefold change of the pressure gain can be attributed to the fact that the velocity distribution on the outer-skull surface has been evened out.

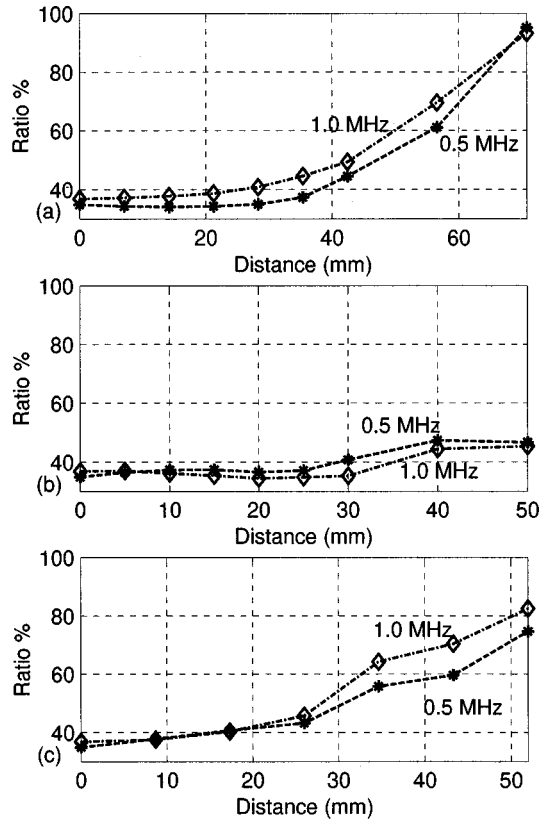


FIG. 13. The percentage of the highest sidelobe pressure amplitude in the x-y plane compared to that at the focal point versus the distance when the focal point is being moved away from the center at both 0.5 and 1.0 MHz: (a) in a horizontal direction; (b) in a vertical direction; (c) in a diagonal direction.

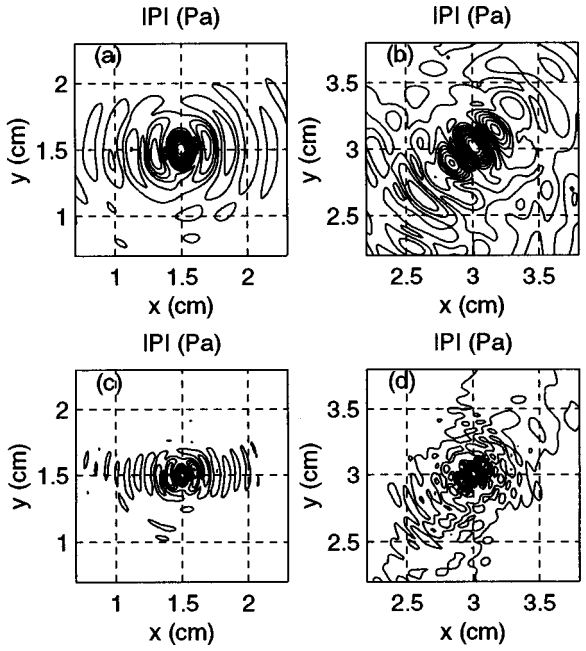


FIG. 14. Example pressure contour lines in the x-y plane when the focus is being moved away from the center in a diagonal direction: (a) and (b) at 0.5 MHz, (c) and (d) at 1.0 MHz; (a) and (c) at (1.5, 1.5, 1.5) cm, (b) and (d) at (3.0, 3.0, 3.0) cm.

It is clear that the focal dimensions along x, y, and z axes are approximately proportional to the acoustic wavelength; this type of linear relationship has been previously pointed out for a slightly curved spherical transducer, e.g., by Fry.[29] Comparing to the results obtained before for the spherical cap transducer (F number is 1.0),[3] the focal dimension along z axis has been significantly reduced; due to the skull (source) geometry, the transverse focal dimensions (focal dimensions along x and y axes) have also been reduced; hence, the overall aspect ratios of the focal volume are more favorable. To a certain extent, the aspect ratios of the focal volume can be controlled by varying the skull surface penetration (radiating) area, which added a new dimension of freedom. With the increased skull penetration (radiating) area, it is clear that a smaller focal volume can be achieved with a given frequency, i.e., a lower frequency can now be used to treat a small target volume.

The limitations of the current study are: the model was for liquid layers; no multiple reflections of standing waves inside the skull were considered; no shear waves in the skull bone were considered; and the bone heterogeneities were not considered. Similarly, variations in skull shape and thickness between different human skulls were not investigated.

The current approach should be good enough for the goal of exploring the best-case scenario of transskull ultrasound therapy and surgery, and the results should be considered ideal, i.e., the real gains may be smaller but probably adequate. In the current study, the outer-skull surface was divided into around 80 000 and 300 000 simple sources at 0.5 and 1.0 MHz, respectively, and phase correction was applied to each simple source. From a previous study,[3] it is shown that if one chooses the source element dimension to be around 5–6 times the wavelength, a near-optimum focus can be obtained. With this in mind, if the source element dimension is chosen to be around 5 times the wavelength, the source element number on the outer-skull surface will be around 1% of the total number of the simple sources, i.e., 800 and 3000 elements at 0.5 and 1.0 MHz, respectively. This makes the practical implementation of this technique difficult and expensive, but not impossible. We are currently investigating practical ways of reducing the element number and thus, making the transskull therapy and surgery more feasible.

IV. CONCLUSIONS

The previously developed numerical model[3] has been extended to handle 3-D digitized layer interfaces reconstructed from cross-sectional contours which were obtained from MR scan images. This model is applied to evaluating the transskull ultrasound field in the brain utilizing the maximum available skull surfaces.

By increasing the skull penetration area, it is feasible to avoid the skull overheating during transskull ultrasound therapy and surgery. This is of particular importance in performing transskull ultrasound therapy and surgery using the thermal effect.

The focus can be moved successfully in the deep regions of the brain volume while maintaining adequate pressure and SAR gains. For shallow targets inside a brain, large sidelobes could prevent successful therapy and surgery.

ACKNOWLEDGMENT

This research was supported by PHS Grant No. CA 76558 awarded by NCI/NIH, PHHS.

[1] K. Hynynen and F. Jolesz, ''Demonstration of noninvasive ultrasound brain therapy through an intact skull,'' Ultrasound Med. Biol. **24**(2), 275–283 (1998).
[2] K. Hynynen and J. Sun, ''Trans-skull ultrasound therapy: The feasibility of using image derived skull thickness information to correct the phase distortion,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control (in press).
[3] J. Sun and K. Hynynen, ''Focusing of therapeutic ultrasound through a human skull: A numerical study,'' J. Acoust. Soc. Am. **104**, 1705–1715 (1998).
[4] G. R. ter Haar, ''Ultrasound focal beam surgery,'' Ultrasound Med. Biol. **21**(9), 1089–1100 (1995).
[5] C. R. Hill and G. R. ter Haar, ''Review article: High intensity focused ultrasound-potential for cancer treatment,'' Br. J. Radiol. **68**, 1296–1303 (1995).
[6] W. J. Fry, J. W. Barnard, F. J. Fry, R. F. Krumins, and J. F. Brennan, ''Ultrasonic lessions in the mammalian central nervous system,'' Science **122**, 517–518 (1955).
[7] F. J. Fry, ''Transskull transmission of an intense focused ultrasonic beam,'' Ultrasound Med. Biol. **3**, 179–184 (1977).
[8] F. J. Fry and J. E. Barger, ''Acoustical properties of the human skull,'' J. Acoust. Soc. Am. **63**, 1576–1590 (1977).
[9] F. J. Fry and S. A. Goss, ''Further studies of the transkull transmission of an intense focused ultrasonic beam: Lesion production at 500 kHz,'' Ultrasound Med. Biol. **6**, 33–38 (1980).
[10] F. J. Fry, N. T. Sanghvi, R. F. Morris, S. Smithson, L. Atkinson, K. Dines, T. Franklin, and J. Hastings, ''A focused ultrasound system for tissue volume ablation in deep seated brain sites,'' *Proceedings of IEEE, New York, 1986 Ultrasonics Symposium*, Vol. 2, pp. 1001–1004 (1986).
[11] S. W. Smith, D. J. Phillips, O. T. von Ramm, and F. L. Thurstone, ''Some advances in acoustic imaging through skull,'' in *Symposium on Biological Effects and Characterizations of Ultrasound Sources*, edited by D. G. Hazzard and M. L. Litz (U.S.D.H.E.W. Publication, FDA, Rockville, MD, 1977), pp. 37–52.
[12] S. W. Smith, G. E. Trahey, and O. T. von Ramm, ''Phased array ultrasound imaging through planar tissue layers,'' Ultrasound Med. Biol. **12**, 229–243 (1986).
[13] J.-L. Thomas and M. A. Fink, ''Ultrasonic beam focusing through tissue inhomogeneities with a time reversal mirror: Application to transskull therapy,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 1122–1129 (1996).
[14] M. Tanter, J.-L. Thomas, and M. A. Fink, ''Focusing and steering through absorbing and aberrating layers: Application to ultrasonic propagation through the skull,'' J. Acoust. Soc. Am. **103**, 2403–2410 (1998).
[15] L. Basauri and P. P. Lele, ''A simple method for production of trackless focal lessions with focused ultrasound: Statistical evaluation of the effects of irradiation on the central nervous system of the cat,'' J. Physiol. (London) **160**, 513–534 (1962).
[16] F. Dunn and F. Fry, ''Ultrasonic threshold dosages for the mammalian central nervous system,'' IEEE Trans. Biomed. Eng. **18**, 253–256 (1971).
[17] P. P. Lele and A. D. Pierce, ''The thermal hypothesis of the mechanism of ultrasonic focal destruction in organized tissues,'' *Interactions of Ultrasound and Biological Tissues—Workshop Proceedings*, edited by J. M. Reid and M. R. Sikov (U.S.D.H.E.W. Publication, FDA, 73-8008 BRH/DBE, 1972).
[18] T. C. Lele and P. P. Lele, ''An analysis of lesion development in the brain and inplastics by high-intensity focussed ultrasound at low-megahertz frequencies,'' J. Acoust. Soc. Am. **51**, 1333–1351 (1972).
[19] N. I. Vykhodtseva, K. Hynynen, and C. Damianou, ''Histologic effects of high intensity pulsed ultrasound exposure with subharmonic emission in rabbit brain *in vivo*,'' Ultrasound Med. Biol. **21(7)**, 969–979 (1995).
[20] N. I. Vykhodtseva, K. Hynynen, and C. Damianou, ''Pulse duration and peak intensity during focused ultrasound surgery: Theoretical and experimental effects in rabbit brain *in vivo*,'' Ultrasound Med. Biol. **20**(9), 987–1000 (1994).
[21] X. Fan and K. Hynynen, ''The effects of wave reflection and refraction at soft tissue interfaces during ultrasound hyperthermia treatments,'' J. Acoust. Soc. Am. **91**, 1727–1736 (1992).
[22] X. Fan and K. Hynynen, ''The effects of curved tissue layers on the power deposition patterns of therapeutic ultrasound beams,'' Med. Phys. **21**, 25–34 (1994).
[23] X. Fan and K. Hynynen, ''Control of the necrosed tissue volume during noninvasive ultrasound surgery using a 16-element phased array,'' Med. Phys. **22**, 297–306 (1995).
[24] E. Kühnicke, ''Calculation of three-dimensional harmonic waves in layered media,'' *Proceedings of 1995 IEEE, New York, Ultrasonics Symposium*, pp. 811–815 (1995).
[25] E. Kühnicke, ''Simulation calculation for monofrequent sound fields in layered media,'' in *Acoustical Imaging*, edited by J. P. Jones (Plenum, New York, 1995), Vol. 21, pp. 47–53.
[26] E. Kühnicke, ''Directional field of a point source for calculation of three-dimensional harmonic waves in layered media,'' in *Acoustical Imaging*, edited by P. Tortoli (Plenum, New York, 1996), Vol. 22.
[27] E. Kühnicke, ''Three-dimensional waves in layered media with nonparallel and curved interfaces: A theoretical approach,'' J. Acoust. Soc. Am. **100**, 709–716 (1996).
[28] M. C. Junger and D. Feit, *Sound, Structures, and Their Interaction* (Acoustical Society of America, Woodbury, NY, 1993).
[29] F. J. Fry, ''Intense focused ultrasound in medicine,'' European Urol. **23**, 2–7 (1993).

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.

Ultrasonic surface waves above a doubly periodic grating

J. F. Allard
Institut d'Acoustique et de Mécanique, Université du Maine UMR CNRS 6613, Avenue Olivier Messiaen 72085, Le Mans Cedex 9, France

L. Kelders and W. Lauriks[a)]
Laboratorium voor Akoestiek en Thermische Fysica, Katholieke Universiteit Leuven, Celestijnenlaan 200D, 3001 Heverlee, Belgium

(Received 18 May 1998; revised 4 December 1998; accepted 4 December 1998)

Ultrasonic surface waves above a layer indented by a doubly periodic set of holes having a square-shaped cross section are studied experimentally. Measurements are compared with predictions obtained from a modal model inspired by the theory of the diffraction of electromagnetic waves. Predictions obtained with the modal model and a simplified model where the periodic structure is replaced by an impedance plane are compared. © *1999 Acoustical Society of America.* [S0001-4966(99)04603-2]

PACS numbers: 43.35.Pt [HEB]

INTRODUCTION

Models for the description of the acoustic surface waves over rough hard periodic and stochastic surfaces have been developed by Tolstoy.[1] The problem of sound propagation over these surfaces has possible applications in aeroacoustics and ocean acoustics. The occurrence of surface waves of large amplitude over honeycomb structures, like those used for sound absorption in aircraft engines, is also a related problem. Two periodic structures are presented in Fig. 1, a rectangular-groove grating having a periodicity in the x direction, and a plate indented with square holes, with a double periodicity in the x and y directions. Let d be the spatial period, a be the thickness of the grooves and the side of the square holes, and h be the depth of the grooves and of the square holes. In Ref. 1, the periodic structures are replaced by an impedance plane Z at $z=0$. The surface wave can be considered as a reflected wave related to a pole of the reflection coefficient R of this impedance plane. Let Z_c and k be the characteristic impedance and the wave number in air, and k_z the z component of the wave number vector of the surface wave. The reflection coefficient can be written as

$$R=\frac{Z+Z_c k/k_z}{Z-Z_c k/k_z}, \tag{1}$$

and k_z is equal to $Z_c k/Z$. The complex diffraction problem is replaced by a simpler reflection problem. In Ref. 1, Tolstoy suggests a method for the calculation of Z from the geometry of the hard surface. A simplified version of Tolstoy's model has been developed by Attenborough[2] for the case of rough surfaces with a finite impedance. A formal analogy exists between the diffraction of electromagnetic waves by conducting surfaces and of acoustic waves by hard surfaces.[3,4] Maystre[5] has pointed out that a precise description of diffraction of electromagnetic waves is generally a difficult and complex work. For some geometries like rectangular or triangular[6] grooves, modal models can be used.[6,7] Using these models in acoustics, it has been shown that for triangular groove gratings, the Tolstoy model does not provide an accurate description of the surface wave. For rectangular groove gratings precise predictions can be obtained if Z in Eq. (1) is replaced by the impedance at normal incidence of a porous material having the same porosity $\phi=a/d$ as the corrugated surface, and a depth h' for the grooves smaller than the actual depth h: $h'=h-a/\pi \log 2$. A similar correction exists for TM waves.[8]

The wave number inside the grooves is equal to k if losses are neglected. Then Z in Eq (1) is given by $Z=-i/\phi Z_c \cot kh'$ and the z wave number component of the surface wave is given by $k_z=i\phi k \tan(kh')$. The damping coefficient in the z direction is $k\phi \tan(kh')$, and the phase velocity in the x direction $c_x=\omega/k_x$ is smaller than the sound speed in free air c_0: $c_x=c_0/(1+\phi^2 \tan^2(kh'))^{1/2}$. The surface wave can be considered as a reflected wave related to a complex angle of incidence θ such that $\sin\theta=(1+\phi^2 \times \tan^2 kh')^{1/2}>1$ and $\cos\theta=i\phi \tan kh'$. Previous experiments performed with honeycomb structures[9] show that these equations can be used to describe accurately the surface wave, a being the distance between two parallel faces of the honeycomb. The complexity of the modes in a honeycomb prevent

a)Electronic mail: Walter.Lauriks@fys.kuleuven.ac.be

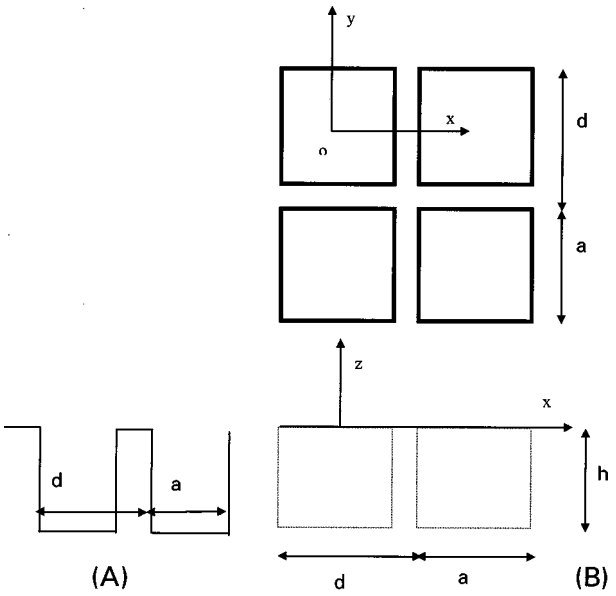


FIG. 1. Rectangular groove grating (a) and a double periodic structure (b).

justification of these equations by a modal model. A study of the surface wave over a simpler doubly periodic grating, the plate indented with square holes [see Fig. 1(b)] is performed in what follows.

I. THE MODAL MODEL

The model presented is a generalization of a previous study concerning ultrasonic surface waves above rectangular groove gratings with a method developed by Hessel *et al.*[7] to describe the diffraction of electromagnetic waves by similar surfaces. The validity of the model is restricted to the case where the wavelength λ is much larger than the period d. The diffraction of electromagnetic waves by doubly periodic structures has been studied by McPhedran *et al.*[10] As indicated in Ref. 10, an incident plane wave is related to an infinite set of diffracted waves specified by a pair of integers (p,q). The components of the wave number vectors of the incident wave are denoted by $\alpha=k_x$, $\beta=k_y$, $\gamma=k_z$ and are related by $\alpha^2+\beta^2+\gamma^2=k^2$. These components for the diffracted waves are

$$\alpha_{p,q}=\alpha+2\pi p/d, \tag{2}$$

$$\beta_{p,q}=\beta+2\pi q/d, \tag{3}$$

$$\gamma_{p,q}=\pm(k^2-\alpha_{p,q}^2-\beta_{p,q}^2)^{1/2}. \tag{4}$$

The sign for $\gamma_{p,q}$ must be chosen so that $\mathrm{Im}(\gamma_{p,q})>0$ and $\gamma_{0,0}=-\gamma$.

A true surface wave over a plane surface is related to a pole of the reflection coefficient. Similarly, for the case of a grating, the diffracted waves are a set of surface waves when their amplitudes are solutions of a singular system of equations, i.e., these amplitudes can reach finite values with an amplitude of the incident wave equal to zero. Previous studies[3,4] of the surface waves over simple gratings have been performed, the losses in the gratings being neglected. The propagation axes of the surface waves are parallel to the xy plane, and the z wave number component of these waves is imaginary. Transposing these results in the present study implies that γ_0 is imaginary, α and β are reals, and $\alpha^2+\beta^2>k^2$. Then Eq. (4) can be rewritten as $\gamma_{p,q}=i((\alpha+2\pi p/d)^2+(\beta+2\pi q/d)^2-k^2)^{1/2}$. It may be noticed that in the previous studies,[3,4] the reflected wave, which corresponds to the zeroth order of diffraction, is the one surface wave which can be detected, the higher orders being too damped in the z direction. The same property will hold for bigratings when $\lambda\gg d$. Let $\hat{v}_i$ and $\hat{v}_{p,q}$ be the components of the velocity fields for the incident and the refracted waves in the plane $z=0$. They can be written $\hat{v}_i=\hat{V}_i\hat{e}_i$ and $\hat{v}_{p,q}=\hat{V}_{p,q}\hat{e}_{p,q}$, where $\hat{e}_i$ and $\hat{e}_{p,q}$ are given by

$$\hat{e}_i=\exp(i(\alpha x+\beta y)), \tag{5}$$

$$\hat{e}_{p,q}=\exp(i((\alpha+2\pi p/d)x+(\beta+2\pi q/d)y)). \tag{6}$$

The pressure fields at $z=0$ are given by

$$\hat{p}_i=\hat{Z}_i\hat{v}_i \tag{7}$$

and

$$\hat{p}_{p,q}=\hat{Z}_{p,q}\hat{v}_{p,q}, \tag{8}$$

where

$$\hat{Z}_i=iZ_ck/(k^2-\alpha^2-\beta^2)^{1/2}, \tag{9}$$

$$\hat{Z}_{p,q}=-iZ_ck/(k^2-(\alpha+2\pi p/d)^2-(\beta+2\pi q/d)^2)^{1/2}. \tag{10}$$

In a hole, the acoustic field is a superposition of different modes, and the z velocity component at $z=0$ is given by $v_{n,m}=V_{n,m}e_{n,m}$, $n=0,1,2,\ldots$; $m=0,1,2,\ldots$. For the hole centered at 0, $e_{n,m}$ can be rewritten

$$e_{n,m}=\cos\left(\frac{n\pi}{a}\left(x+\frac{a}{2}\right)\right)\cos\left(\frac{m\pi}{a}\left(y+\frac{a}{2}\right)\right), \tag{11}$$

and the related pressure field $p_{n,m}$ is given by $p_{n,m}=Z_{n,m}v_{n,m}$, where

$$Z_{n,m}=-iZ_c\frac{k}{k_{n,m}}\cot(k_{n,m}h). \tag{12}$$

The z wave number vector component $k_{n,m}$ is given by $k_{n,m}=\pm i((n^2+m^2)\pi^2/a^2-k^2)^{1/2}$. The continuity conditions in the $z=0$ plane for the z component of the velocity are

$$\hat{v}_i+\sum_{p,q}\hat{v}_{p,q}=\sum_{n,m}v_{n,m},\quad |x|<\frac{a}{2},\quad |y|<\frac{a}{2}, \tag{13a}$$

$$\hat{v}_i+\sum_{p,q}\hat{v}_{p,q}=0,\quad \frac{d}{2}>|x|>\frac{a}{2},\quad \frac{d}{2}>|y|>\frac{a}{2}. \tag{13b}$$

For the pressure, the continuity equation is

$$\hat{p}_i+\sum_{p,q}\hat{p}_{p,q}=\sum_{n,m}p_{n,m},\quad |x|<\frac{a}{2},\quad |y|<\frac{a}{2}. \tag{14}$$

Equations (13) and (14) can be replaced by a set of linear equations:

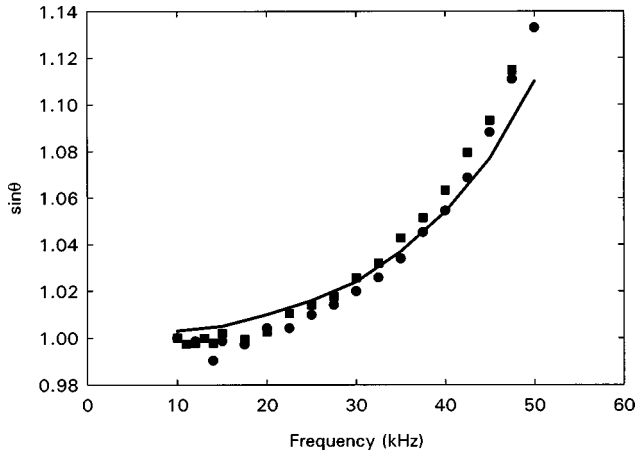


FIG. 2. Measured and predicted sin θ for the doubly periodic structure. Measurement in the direction x and y ●●●. Measurement in the direction parallel to the diagonals of the squares ■■■. Predictions ———.

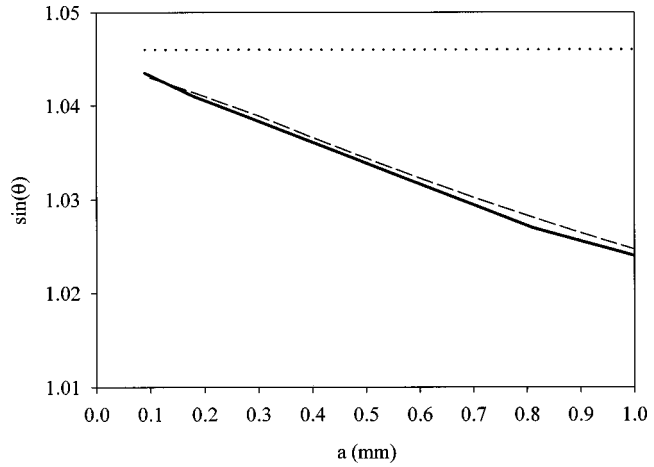


FIG. 3. Evaluation of sin θ for $a/d=1/1.33$, $h=0.9$ mm at 30 kHz. Modal model ———; simple model using corrected depth h' ---; simple model using actual depth $h\cdots$.

$$\sum_{p,q}\left[\sum_{n',m'} V_{n',m'}\langle e_{n',m'},\hat{e}^{*}_{p,q}\rangle\langle\hat{e}_{p,q},e_{n,m}\rangle\right]\frac{\hat{Z}_{p,q}}{d^2} - V_{n,m}Z_{n,m}\frac{a^2}{4}(1+\delta_{n,0})(1+\delta_{m,0}) = -\hat{V}_i\hat{Z}_i\langle\hat{e}_i,e_{n,m}\rangle. \quad (15)$$

In these equations $\langle f,g\rangle$ is defined as $\langle f,g\rangle = \int_{-a/2}^{a/2}\int_{-a/2}^{a/2} fg\,dx\,dy$ and δ is the Kronecker symbol. This system of equations is truncated by using a finite number of space harmonics ($|p|\leqslant 19$, $|q|\leqslant 19$) and a finite number of modes in the holes ($n\leqslant 5$, $m\leqslant 5$). The truncated system of 25 equations with 25 unknowns $V_{n,m}$ is singular when the determinant of the system is equal to zero. It has been verified that the results presented in the next section are not noticeably different with a more drastic truncation ($|p|\leqslant 15$, $|q|\leqslant 15$, $n\leqslant 4$, $m\leqslant 4$).

II. COMPARISON BETWEEN MEASUREMENT AND PREDICTION

A doubly periodic structure similar to the one represented in Fig. 1(b) has been used. The dimensions of the holes are $h=0.9$ mm, $a=1.0$ mm, and the spatial period $d=1.33$ mm. Measurements of the phase velocity c of the surface wave have been performed in the directions parallel to the sides and the diagonal of the squares. An oscillator drives a PANAMETRICS V3052 piezoelectric transducer. A Bruël and Kjær 4138 high-frequency condenser microphone is used as detector. The phase velocity has been measured by recording the phase of a sine signal as well as the time of flight of a burst. The quantity c_0/c is sin θ and can be written as $c_0/c=\sin\theta=(\alpha^2+\beta^2)^{1/2}/k$. This quantity is represented in Fig. 2 for both directions. The measured values are close to each other for both directions. The predicted sin θ for the direction parallel to x is obtained by setting $\alpha=k\sin\theta$, $\beta=0$. A simple program is used to obtain sin θ such as the determinant of the system of the 25 equations [Eq. (15)] is equal to zero. For the direction parallel to the diagonal, the predicted sinθ is obtained by setting $\alpha=\beta=k\sin\theta/\sqrt{2}$. The predicted values are very close to each other for both directions and are represented on the same curve. There is a good agreement between measurement and prediction.

III. SIMPLE MODEL

If only the mode $n=0$, $m=0$ is taken into account, the acoustic surface impedance of a hole, i.e., the ratio p/v_z at $z=0$ in the hole, is equal to $-iZ_c\cot(kh)$. The classical models used for porous media lead to an averaged normal velocity over the surface given by $\hat{v}_z=v_z\phi$, where ϕ is the porosity [for the bigrating $\phi=(a/d)^2$] and a surface impedance is equal to $Z=-i/\phi\,Z_c\cot(kh)$. It has been shown that for the Tolstoy model[1] (for rectangular groove gratings) the surface impedance can be used in Eq. (1) with a corrected depth,[3] i.e., h must be replaced by h'. This model can be transposed directly to the case of the doubly periodic structure. The same formulas, with the definitions of the symbols a, d, h, related to Fig. 1(b), can be used. In order to test the validity of this simple model, the predicted value of sinθ with the modal model at 30 kHz for $h=0.9$ mm and $a/d=1/1.33$ is represented as a function of a, and compared with this simple model. There is a good agreement between the modal model and this simple model. Predictions obtained with this simple model with the actual depth instead of h' are also shown in Fig. 3. It appears that using the surface impedance Z in Eq. (1) with the actual depth h gives predictions which can be very different from those obtained with the modal model.

IV. CONCLUSION

The modal method provides precise predictions of the wave number vector of the surface waves over a plane perforated by square holes. The predictions obtained in a simpler way by replacing the diffracting grating by a reflecting locally reacting porous material are close to those obtained by the modal model only if the lateral dimensions of the holes are much smaller than the depth.

ACKNOWLEDGMENT

The authors thank the Plan Pluriformation Surface (Le Mans) for its financial support.

[1] I. Tolstoy, ''Smoothed boundary conditions, coherent low-frequency scatter and boundary modes,'' J. Acoust. Soc. Am. **75**, 1–22 (1984).
[2] K. Attenborough, ''Propagation from a point source over a rough finite impedance boundary,'' J. Acoust. Soc. Am. **98**, 1717–1722 (1995).
[3] L. Kelders, J. F. Allard, and W. Lauriks, ''Ultrasonic surface waves above rectangular-groove gratings,'' J. Acoust. Soc. Am. **103**, 2730–2733 (1998).
[4] W. Lauriks, L. Kelders, and J. F. Allard, ''Surface waves above a grating having a triangular profile,'' Ultrasonics **36**, 865–871 (1998).
[5] D. Maystre, ''General study of grating anomalies from electromagnetic surface modes,'' in *Electromagnetic Surface Modes*, edited by A. Boardman (Wiley, New York, 1982).
[6] S. Jovicevic and S. Sesnic, ''Diffraction of a parallel and perpendicular-polarised wave from an echelette grating,'' J. Opt. Soc. Am. **62**, 865–877 (1972).
[7] A. Hessel, J. Schmoys, and D. Y. Tseng, ''Bragg-angle blazing of diffraction gratings,'' J. Opt. Soc. Am. **65**, 380–384 (1975).
[8] R. A. Hurd, ''The propagation of an electromagnetic wave along an infinite corrugated surface,'' Can. J. Phys. **32**, 727–734 (1954).
[9] J. Tizianel, J. F. Allard, and B. Brouard, ''Surface waves above honeycombs,'' J. Acoust. Soc. Am. **104**, 2525–2528 (1998).
[10] C. McPhedran, G. H. Derrick, and L. C. Botten, ''Theory of crossed gratings,'' in *Electromagnetic Theory of Gratings*, edited by R. Petit (Springer, New York, 1980).

The estimation of signal-to-noise ratio in continuous speech for disordered voices

Yingyong Qi
Department of Speech and Hearing Sciences and Department of Electrical and Computer Engineering, University of Arizona, Tucson, Arizona 85721

Robert E. Hillman
Voice and Speech Laboratory, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts, Department of Otology and Laryngology, Harvard Medical School, Communication Science and Disorders, MGH-Institute of Health Professions, and Research Laboratory of Electronics, Massachusetts Institute of Technology

Claudio Milstein
Voice and Speech Laboratory, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts and Department of Speech and Hearing Sciences, University of Arizona, Tucson, Arizona 85721

(Received 3 November 1998; revised 29 December 1998; accepted 11 January 1999)

Presented is a method of estimating the signal-to-noise ratio (SNR) of continuous utterances for patients with various types of voice disorders that ranged in severity of dysphonia from mild to severe. The SNR is estimated based on the residual that is left after systematically removing the short- and long-term correlations that exist in the speech signal. Results indicate that the SNR is consistent with human perceptual judgments, particularly those that consistently differentiate close-to-normal versus highly disphonic voices. © *1999 Acoustical Society of America.* [S0001-4966(99)01604-5]

PACS numbers: 43.70.Gr [AL]

INTRODUCTION

In current research and clinical practice, acoustic analyses of speech signals often rely on vowel phonations that are sustained for several seconds. A number of methods have been developed to analyze a selected segments from such samples of sustained vowel phonation. These include perturbations of fundamental frequency and amplitude, and harmonics-to-noise ratio (Horii, 1980; Yumoto *et al.*, 1982; Qi and Hillman, 1997)

Sustained vowel phonation, however, is not necessarily a valid representation of an individual's vocal function during continuous speech. For example, real running speech involves constant and rapid adjustments of vocal mechanisms (e.g., rapid initiation and termination of voicing) that are not present during sustained phonation of a vowel. Thus, it would be desirable, and potentially more valid, to obtain estimates from continuous speech of acoustic parameters that are associated with abnormalities in voice quality. Toward this goal, we here introduce a method for acoustically estimating the signal-to-noise ratio (SNR) of continuous utterances, which we hope will be useful both for clinical and research-related evaluations of voice production. This method is evaluated by comparing its results with human perceptual evaluations.

I. METHODS

In the analysis proposed here, speech signals are decomposed into two components: a correlated/predictable component (signal) and an uncorrelated/unpredictable component (noise). The SNR, thus, defines the strength of the correlated component of a speech signal relative to the uncorrelated/unpredictable, noise component. The decomposition of the speech signal into correlated and uncorrelated components is accomplished by systematically removing existing correlations from the signal until the residual signal appears to be a random Gaussian (normally distributed) sequence (Schroeder and Atal, 1985). This approach is similar to the statistical procedures used in analysis of variance, where known variables are successively factored out until the remaining variations appear to be random with a normal distribution.

According to the acoustic theory of speech production, there are two known types of correlations present in a speech signal: a short-term correlation and a long-term correlation (Schroeder and Atal, 1985). Short-term correlation refers to the correlation/predictability of a signal on a sample-by-sample basis. Such a correlation is primarily associated with the resonances of the vocal tract. For example, the pattern of oscillation (resonance) within each fundamental period is predictable, i.e., the magnitude of the current sample could be predicted from the samples that immediately precede the current sample when the formants of vocal tract are known (Fant, 1981). This short-term predictability would be disrupted by the glottal input for the next cycle and/or by any random variations.

Long-term correlation refers to the correlation/predictability of the signal based on samples that do not immediately precede the current sample. Such a correlation is primarily associated with the quasi-periodical nature of voice production (Ramachandran and Kabal, 1989). For example, the signal characteristics around the beginning of each cycle would be predictable, to a certain extent, based on information from around the beginning of previous cycles. This pre-

diction would be disrupted by the onset/offset of voicing and/or by any random variations.

The decomposition of speech signals into short- and long-term correlations plus Gaussian noise has been successfully applied in telecommunication systems (Schroeder and Atal, 1985). In the code-excited linear prediction (CELP) based speech coders of some cellular phone systems, for example, only parameters related to the short- and long-term correlations are transmitted. Speech signals are reconstructed at the receiver by adding (filtering) random Gaussian noise using the transmitted short- and long-term correlation coefficients. Although it is necessary to synthesize a random noise in the receiver that has similar variance and temporal distribution as that in the transmitter, the unpredictable, noise component of speech is **not** transmitted in the cellular system. The adequacy of decomposing speech signals into short- and long-term correlated components plus a noise component is demonstrated by the fact that cellular phones provide adequate speech quality for normal communication.

A number of predictions could be made about the decomposition of speech signals. The residual signal, for example, should approximate a Gaussian process. This has been well demonstrated in previous publications (Schroeder and Atal, 1985). In this work, the proposed SNR was evaluated by comparing it to human perceptual ratings of a relatively large set of speech samples.

A. Subjects and recordings

Speech samples were recorded at the Voice and Speech Laboratory of Massachusetts Eye and Ear Infirmary. Eighty-seven subjects (40 men and 47 women) diagnosed with a wide variety of laryngeal voice pathologies provided the speech samples. Each subject was asked to read the Rainbow Passage at comfortable fundamental frequency and intensity levels. Audio recordings were made using a condenser microphone (Sennheiser) and a digital tape recorder (Tascom, DA-30) in a sound treated booth. The microphone was suspended a constant distance of 15 cm from the lips of each subject using a head-mounted device. All recordings were low-pass filtered ($f_c = 7.5$ kHz) and redigitized into a computer at a sampling rate of 16 kHz and a 16-bit A/D resolution. The first two sentences of the Rainbow Passage were used for subsequent acoustic analysis and perceptual evaluation.

B. Acoustic analysis

Linear prediction (LP) was used to determine both short- and long-term correlations (Markel and Gray, 1976; Ramachandran and Kabal, 1989). For short-term correlation, LP analysis was made on a window-by-window (no overlap) basis. The LP filter was obtained using a Hamming window with window length of 20 ms. The order of the LP filter was 14 (Markel and Gray, 1976). To remove short-term correlation, the original signal was inverse filtered by the LP filter using overlap save to ensure continuity during filter update. The residual signal of this LP inverse filtering was the short-term, decorrelated signal which was then further processed for long-term decorrelation.

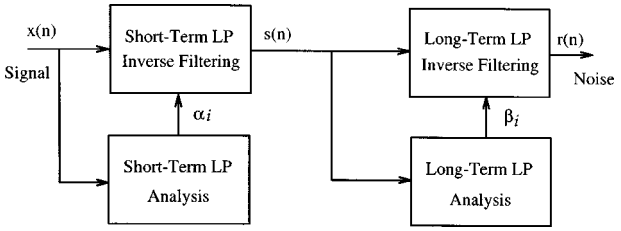


FIG. 1. Flow chart of the short- and long-term decorrelation process.

During long-term decorrelation, linear prediction was made based on samples that were not immediately preceding the current sample of the short-term decorrelated, residue signal. The window length for minimizing prediction error was 2.5 ms, which is long enough to include the pulselike peaks of short-term LP residual signals that often occur during voiced segments of speech. Because the exact location of the next residual peak varies somewhat from cycle to cycle, the closest sample used for making prediction was between 1.25 to 17.5 ms prior to the first sample to be predicted. Thus, this included the fundamental frequency range from 60–800 Hz in the predictive analysis. The LP filter that produced the minimal prediction error over this sliding range was chosen as the final long-term LP filter. The order of the filter was 3 (Ramachandran and Kabal, 1989). To remove the long-term correlation, the short-term decorrelated residual signal was inverse filtered by the long-term LP filter. Overlap save was used again to ensure continuity during filter update. The output of this second stage of inverse filtering was considered to be the final (short- and long-term decorrelated) noise component of the speech signal. A flow chart of the short- and long-term decorrelation processes is shown in Fig. 1. Example signals are shown in Fig. 2.

The final SNR was computed as the ratio of average rms amplitude between the original signal and its corresponding short- and long-term decorrelated signal. This ratio was reduced by one before converting it to dB scale because the original signal represents signal plus noise.

C. Perceptual evaluations

The recorded voice samples (87 in total) were perceptually rated by the same group of listeners using two different types of scales: a categorical scale and a continuous scale. The two ratings were made about three months apart to minimize any potential learning effects. Judges consisted of five speech pathologists with normal hearing (screened at 25 dB for speech frequencies) and extensive training and experience in the diagnosis and treatment of voice disorders. All ratings were accomplished using an interactive graphical user interface on a computer with stimuli presented over headphone. Stimuli consisted of the first two sentences of the Rainbow Passage.

In the categorical rating task, judges were asked to classify each voice sample as (1) normal, (2) mild, (3) mild-to-moderate, (4) moderate, (5) moderate to severe, (6) severe, or (7) aphonic. The judges were allowed to listen to each voice sample as many times as they wished before entering their response. For assessing intrajudge reliability, each

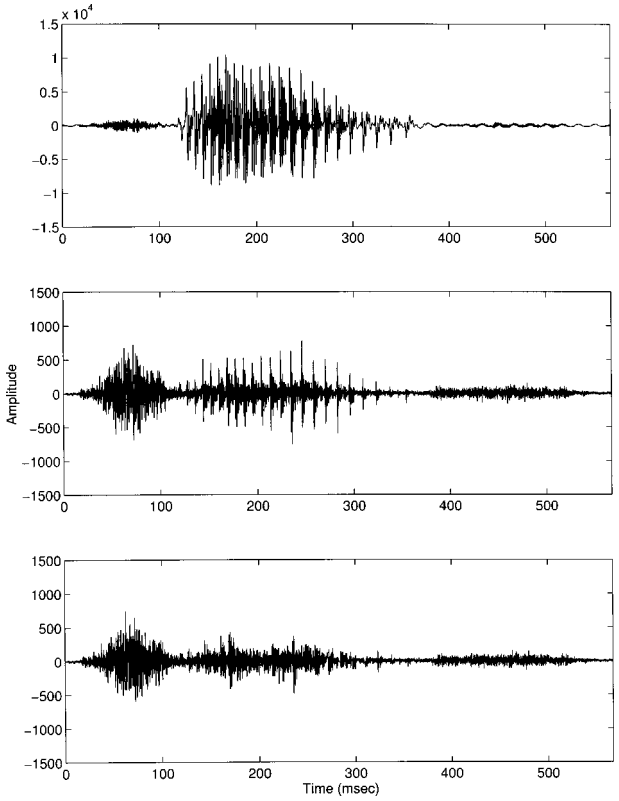


FIG. 2. Example of original signal (top), short-term decorrelated signal (middle), and short- and long-term decorrelated signal (bottom). The original signal was recorded from a male talker saying the word ''choice.''

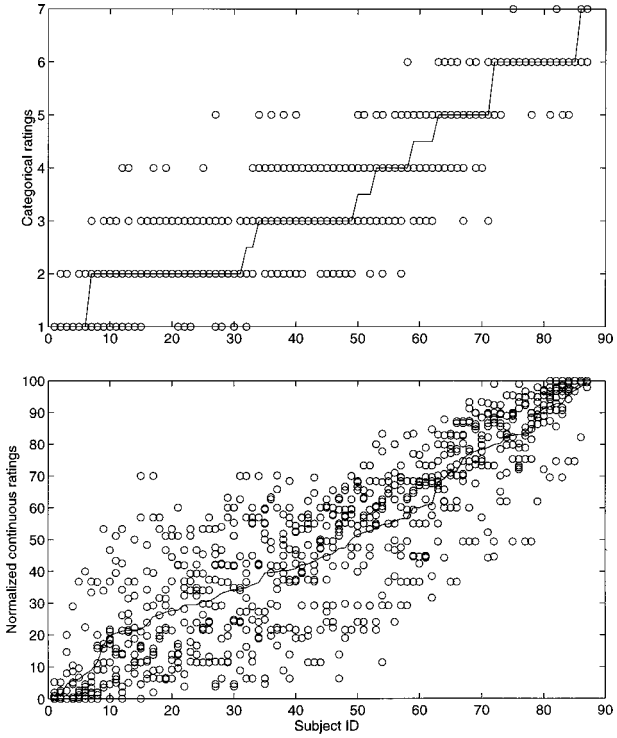


FIG. 3. Categorical ratings for all judges and voice samples with associated median values (top) and normalized, continuous ratings for all judges and voice samples with associated mean values (bottom).

judge repeated the entire rating session twice for a different random ordering of stimulus presentation with a period of at least 24 h in between rating sessions.

In the continuous rating task, judges were asked to use a number to describe the degree of perceived dysphonia relative to a standard voice sample. The standard voice sample was assigned a number of 100. Voices perceived to have more dysphonia than the standard, for example, would be given a rating of more than 100, and voices with less dysphonia, would be assigned a number less than 100. Raters were free to assign any value, as high or as low as they considered necessary. They also had access to the reference sample at all times, and were allowed to listen to each voice sample as many times as they wished before entering their response. As was the case for the categorical task, intrajudge reliability was assessed by repeating the entire rating session (stimuli in different random order) on a different day.

II. RESULTS AND CONCLUSIONS

The categorical ratings and their medians are shown in Fig. 3 (top) for each voice sample. Spearman correlations were computed for all ten judgments (5 judges$\times$2 sessions). Results indicated that all correlations were significant ($p<0.001$). Intrajudge correlations ranged from 0.87 to 0.93 and interjudge correlations ranged from 0.82 to 0.91.

The continuous ratings and their means are shown in Fig. 3 (bottom) for each voice sample. Here, all scores were normalized to the range of 0–100 based on the maximum and minimum scores of a rating session. Pearson correlations were computed for all ten judgments (5 judges$\times$2 sessions). Results indicate that all correlations were significant ($p<0.001$). Intrajudge correlations ranged from 0.93 to 0.96. Interjudge reliability was calculated using Cronbach's α (Cronbach, 1970). This statistics entails measuring the correlation between each individual listener's mean rating for each stimulus with the group mean of all the other listeners. Cronbach's α was 0.97, indicating adequate reliability among listeners in the continuous scaling task.

As shown, despite significantly high intra- and interjudge correlations, both categorical and continuous ratings extend (overlap) over a relatively large range for most voice samples. For example, a voice sample in the middle of the perceptual scale could have a categorical rating ranging from mild to severe or a continuous rating ranging from 25 to 75. This overlap, however, is minimal between samples that are rated as close-to-normal (normal and mild, 31 samples) and those that are rated as highly disphonic (moderate to severe, severe, and aphonic, 25 samples). A histogram of the categorical scores for these two subgroups ($31+25=57$ total samples) is shown in Fig. 4. As expected, the judges appeared inconsistent when viewing their performance across the entire set of voice samples, but they were able to differentiate between close-to-normal and highly disphonic voice samples quite well.

The median categorical ratings and mean continuous ratings are shown in Fig. 5 as a function of the computed SNRs for all subjects. The Spearman correlation between the categorical ratings and SNRs ($r=-0.76$, $p<0.001$) and the Pearson correlation between the continuous ratings and SNRs ($r=-0.78$, $p<0.001$) are both statistically significant, but relatively low in terms of the amount of variation

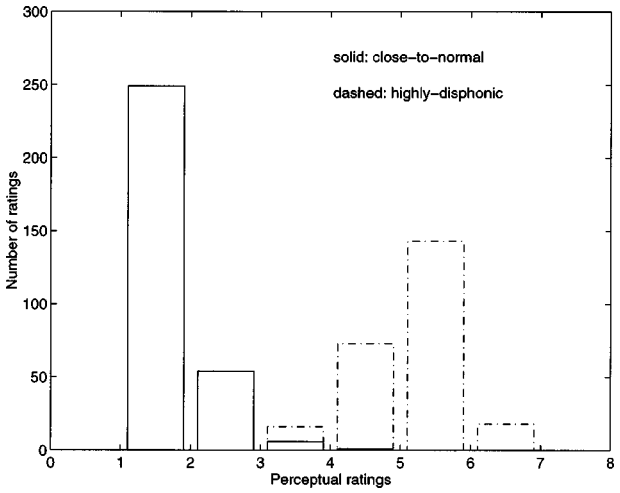


FIG. 4. Histogram for voice samples whose perceptual ratings fell in either the close-to-normal or highly disphonic range.

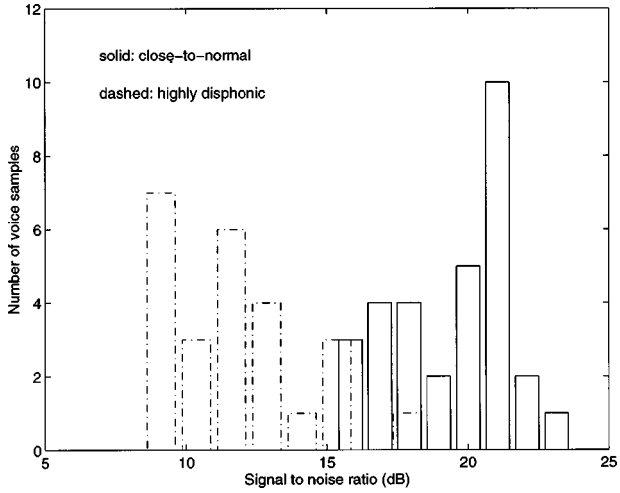


FIG. 6. Histogram of SNR values for voice samples whose perceptual ratings are either in the close-to-normal or highly disphonic range.

actually accounted for ($r^2 < 0.61$). By way of comparison, the SNRs for samples that are rated as close-to-normal or highly disphonic are shown in Fig. 6. There is a clear separation (54/57=95%) in SNR between these two groups, indicating that the computed SNRs are in agreement with the perceptual ratings when the ratings are made consistently.

These results seem to indicate that the proposed SNR, similar to the harmonics-to-noise ratio for vowels, has a moderate degree of correlation with perceptual ratings of human listeners (Yumoto *et al.*, 1982). Obviously, experiments undertaken here merely provide some preliminary support for the proposed SNR measurement. Further experiments are necessary to more rigorously establish the relationship between SNR and specific aspects of pathological voice/speech production/perception. To date, comprehensive understanding and agreement on methods for evaluating pathological voice production/perception is lacking. The SNR measurement described here is developed largely based on estimations of short- and long-term correlations that have been successfully applied to modern telecommunication systems. It represents a first attempt to directly quantify acoustic properties of continuous utterance for disordered voices. It is our hope that the proposed SNR measure could be developed into a useful tool for clinical and research-related voice assessment.

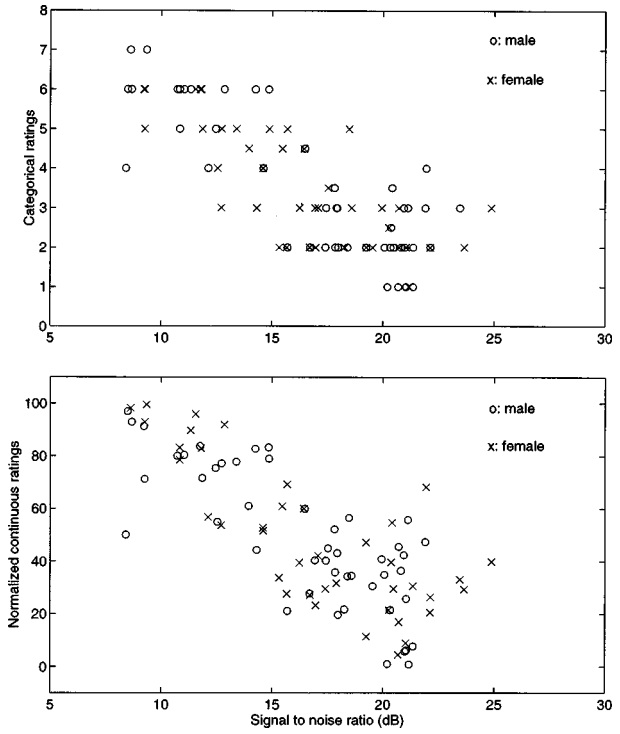


FIG. 5. Categorical (top) and continuous (bottom) ratings as a function of the computed SNR.

ACKNOWLEDGMENT

This work was supported, in part, by grants from the National Institute of Deafness and Other Communication Disorders: DC00266, Objective Assessment of Vocal Hyperfunction.

Cronbach, L. (**1970**). *Essentials of Psychological Testing* (Harper and Row, New York).

Fant, G. (**1981**). ''The source filter concept in voice production,'' STL-QPSR **1**, 21–37.

Horii, Y. (**1980**). ''Vocal shimmer in sustained phonation,'' J. Speech Hear. Res. **23**, 202–209.

Markel, J., and Gray, A. (**1976**). *Linear Prediction of Speech* (Springer-Verlag, Berlin).

Qi, Y., and Hillman, R. (**1997**). ''Temporal and spectral estimations of harmonics-to-noise ratio in human voice signals,'' J. Acoust. Soc. Am. **102**, 537–543.

Ramachandran, R. P., and Kabal, P. (**1989**). ''Pitch prediction filters in speech coding,'' IEEE Trans. Acoust., Speech, Signal Process. **37**, 467–478.

Schroeder, M., and Atal, B. (**1985**). ''Code excited linear prediction (CELP): High quality speech at very low bit rates,'' in Proc. ICASSP, pp. 937–940.

Yumoto, E., Gould, W., and Baer, T. (**1982**). ''Harmonics-to-noise ratio as an index of the degree of hoarseness,'' J. Acoust. Soc. Am. **71**, 1544–1550.